

SEQUENCE LISTING

Sequence No.: 1

Sequence length: 154

Sequence type: Amino acid

Topology: Linear

Sequence kind: Protein

Hypothetical: NO

Original source:

Organism species: *Homo sapiens*

Cell kind: Fibrosarcoma

Cell line: HT-1080

Clone name: HP00658

Sequence description

Met Lys Val Ser Ala Ala Ala Leu Ala Val Ile Leu Ile Ala Thr Ala

1 5 10 15

Leu Cys Ala Pro Ala Ser Ala Ser Pro Tyr Ser Ser Asp Thr Thr Pro

20 25 30

Cys Cys Phe Ala Tyr Ile Ala Arg Pro Leu Pro Arg Ala His Ile Lys

35 40 45

Glu Tyr Phe Tyr Thr Ser Gly Lys Cys Ser Asn Pro Ala Val Val His

50 55 60

Arg Ser Arg Met Pro Lys Arg Glu Gly Gln Gln Val Trp Gln Asp Phe

65 70 75 80

Leu Tyr Asp Ser Arg Leu Asn Lys Gly Lys Leu Cys His Pro Lys Glu

85 90 95

Pro Pro Ser Val Cys Gln Pro Arg Glu Glu Met Gly Ser Gly Val His

100 105 110

65

115 120 125
Thr Gln Ile Cys Leu Phe Leu Leu Ala Leu Val Leu Ala Trp Glu Ala
130 135 140
Ser Pro His Tyr Pro Thr Pro Pro Ala Pro
145 150

Sequence No.: 2

Sequence length: 315

Sequence type: Amino acid

Topology: Linear

Sequence kind: Protein

Hypothetical: No

Original source:

Organism species: *Homo sapiens*

Cell kind: Epidermoid carcinoma

Cell line: KB

Clone name: HP00714

Sequence description

Met Asp Leu Arg Gln Phe Leu Met Cys Leu Ser Leu Cys Thr Ala Phe
1 5 10 15
Ala Leu Ser Lys Pro Thr Glu Lys Lys Asp Arg Val His His Glu Pro
20 25 30
Gln Leu Ser Asp Lys Val His Asn Asp Ala Gln Ser Phe Asp Tyr Asp
35 40 45
His Asp Ala Phe Leu Gly Ala Glu Glu Ala Lys Thr Phe Asp Gln Leu
50 55 60
Thr Pro Glu Glu Ser Lys Glu Arg Leu Gly Lys Ile Val Ser Lys Ile

66

85	90	95
Ile Lys Phe Ala Gln Lys Arg Trp Ile Tyr Glu Asp Val Glu Arg Gln		
100	105	110
Trp Lys Gly His Asp Leu Asn Glu Asp Gly Leu Val Ser Trp Glu Glu		
115	120	125
Tyr Lys Asn Ala Thr Tyr Gly Tyr Val Leu Asp Asp Pro Asp Pro Asp		
130	135	140
Asp Gly Phe Asn Tyr Lys Gln Met Met Val Arg Asp Glu Arg Arg Phe		
145	150	155
Lys Met Ala Asp Lys Asp Gly Asp Leu Ile Ala Thr Lys Glu Glu Phe		
165	170	175
Thr Ala Phe Leu His Pro Glu Glu Tyr Asp Tyr Met Lys Asp Ile Val		
180	185	190
Val Gln Glu Thr Met Glu Asp Ile Asp Lys Asn Ala Asp Gly Phe Ile		
195	200	205
Asp Leu Glu Glu Tyr Ile Gly Asp Met Tyr Ser His Asp Gly Asn Thr		
210	215	220
Asp Glu Pro Glu Trp Val Lys Thr Glu Arg Glu Gln Phe Val Glu Phe		
225	230	235
Arg Asp Lys Asn Arg Asp Gly Lys Met Asp Lys Glu Glu Thr Lys Asp		
245	250	255
Trp Ile Leu Pro Ser Asp Tyr Asp His Ala Glu Ala Glu Ala Arg His		
260	265	270
Leu Val Tyr Glu Ser Asp Gln Asn Lys Asp Gly Lys Leu Thr Lys Glu		
275	280	285
Glu Ile Val Asp Lys Tyr Asp Leu Phe Val Gly Ser Gln Ala Thr Asp		
290	295	300

Sequence No.: 3

Sequence length: 158

Sequence type: Amino acid

Topology: Linear

Sequence kind: Protein

Hypothetical: No

Original source:

Organism species: *Homo sapiens*

Cell kind: Stomach cancer

Clone name: HP00876

Sequence description

Met Ala Ser Arg Ser Met Arg Leu Leu Leu Leu Leu Ser Cys Leu Ala

1 5 10 15

Lys Thr Gly Val Leu Gly Asp Ile Ile Met Arg Pro Ser Cys Ala Pro

20 25 30

Gly Trp Phe Tyr His Lys Ser Asn Cys Tyr Gly Tyr Phe Arg Lys Leu

35 40 45

Arg Asn Trp Ser Asp Ala Glu Leu Glu Cys Gln Ser Tyr Gly Asn Gly

50 55 60

Ala His Leu Ala Ser Ile Leu Ser Leu Lys Glu Ala Ser Thr Ile Ala

65 70 75 80

Glu Tyr Ile Ser Gly Tyr Gln Arg Ser Gln Pro Ile Trp Ile Gly Leu

85 90 95

His Asp Pro Gln Lys Arg Gln Gln Trp Gln Trp Ile Asp Gly Ala Met

100 105 110

Tyr Leu Tyr Arg Ser Trp Ser Gly Lys Ser Met Gly Gly Asn Lys His

130 135 140
Glu Cys Asn Lys Arg Gln His Phe Leu Cys Lys Tyr Arg Pro
145 150 155

Sequence No.: 4

Sequence length: 376

Sequence type: Amino acid

Topology: Linear

Sequence kind: Protein

Hypothetical: No

Original source:

Organism species: *Homo sapiens*

Cell kind: Liver

Clone name: HP01134

Sequence description

Met Val Trp Lys Val Ala Val Phe Leu Ser Val Ala Leu Gly Ile Gly
1 5 10 15
Ala Val Pro Ile Asp Asp Pro Glu Asp Gly Gly Lys His Trp Val Val
20 25 30
Ile Val Ala Gly Ser Asn Gly Trp Tyr Asn Tyr Arg His Gln Ala Asp
35 40 45
Ala Cys His Ala Tyr Gln Ile Ile His Arg Asn Gly Ile Pro Asp Glu
50 55 60
Gln Ile Val Val Met Met Tyr Asp Asp Ile Ala Tyr Ser Glu Asp Asn
65 70 75 80
Pro Thr Pro Gly Ile Val Ile Asn Arg Pro Asn Gly Thr Asp Val Tyr
85 90 95
Gln Gly Val Pro Lys Asp Tyr Thr Gly Glu Asp Val Thr Pro Gln Asn

Phe Leu Ala Val Leu Arg Gly Asp Ala Glu Ala Val Lys Gly Ile Gly
115 120 125
Ser Gly Lys Val Leu Lys Ser Gly Pro Gln Asp His Val Phe Ile Tyr
130 135 140
Phe Thr Asp His Gly Ser Thr Gly Ile Leu Val Phe Pro Asn Glu Asp
145 150 155 160
Leu His Val Lys Asp Leu Asn Glu Thr Ile His Tyr Met Tyr Lys His
165 170 175
Lys Met Tyr Arg Lys Met Val Phe Tyr Ile Glu Ala Cys Glu Ser Gly
180 185 190
Ser Met Met Asn His Leu Pro Asp Asn Ile Asn Val Tyr Ala Thr Thr
195 200 205
Ala Ala Asn Pro Arg Glu Ser Ser Tyr Ala Cys Tyr Tyr Asp Glu Lys
210 215 220
Arg Ser Thr Tyr Leu Gly Asp Trp Tyr Ser Val Asn Trp Met Glu Asp
225 230 235 240
Ser Asp Val Glu Asp Leu Thr Lys Glu Thr Leu His Lys Gln Tyr His
245 250 255
Leu Val Lys Ser His Thr Asn Thr Ser His Val Met Gln Tyr Gly Asn
260 265 270
Lys Thr Ile Ser Thr Met Lys Val Met Gln Phe Gln Gly Met Lys Arg
275 280 285
Lys Ala Ser Ser Pro Val Pro Leu Pro Pro Val Thr His Leu Asp Leu
290 295 300
Thr Pro Ser Pro Asp Val Pro Leu Thr Ile Met Lys Arg Lys Leu Met
305 310 315 320
Asn Thr Asn Asp Leu Glu Glu Ser Arg Gln Leu Thr Glu Glu Ile Gln

70

340 345 350
Asn Leu Cys Glu Lys Pro Tyr Pro Leu His Arg Ile Lys Leu Ser Met
355 360 365
Asp His Val Cys Leu Gly His Tyr
370 375

Sequence No.: 5

Sequence length: 173

Sequence type: Amino acid

Topology: Linear

Sequence kind: Protein

Hypothetical: No

Original source:

Organism species: *Homo sapiens*

Cell kind: Epidermoid carcinoma

Cell line: KB

Clone name: HP10029

Sequence description

Met Ala Ala Pro Ser Gly Gly Trp Asn Gly Val Arg Ala Ser Leu Trp
1 5 10 15
Ala Ala Leu Leu Leu Gly Ala Val Ala Leu Arg Pro Ala Glu Ala Val
20 25 30
Ser Glu Pro Thr Thr Val Ala Phe Asp Val Arg Pro Gly Gly Val Val
35 40 45
His Ser Phe Ser His Asn Val Gly Pro Gly Asp Lys Tyr Thr Cys Met
50 55 60
Phe Thr Tyr Ala Ser Gln Gly Gly Thr Asn Glu Gln Trp Gln Met Ser

71

85

90

95

Pro Gln Gly Lys Ser Tyr Leu Tyr Phe Thr Gln Phe Lys Ala Glu Val

100

105

110

Arg Gly Ala Glu Ile Glu Tyr Ala Met Ala Tyr Ser Lys Ala Ala Phe

115

120

125

Glu Arg Glu Ser Asp Val Pro Leu Lys Thr Glu Glu Phe Glu Val Thr

130

135

140

Lys Thr Ala Val Ala His Arg Pro Gly Ala Phe Lys Ala Glu Leu Ser

145

150

155

160

Lys Leu Val Ile Val Ala Lys Ala Ser Arg Thr Glu Leu

165

170

Sequence No.: 6

Sequence length: 73

Sequence type: Amino acid

Topology: Linear

Sequence kind: Protein

Hypothetical: No

Original source:

Organism species: *Homo sapiens*

Cell kind: Epidermoid carcinoma

Cell line: KB

Clone name: HP10189

Sequence description

Met Gly Val Lys Leu Glu Ile Phe Arg Met Ile Ile Tyr Leu Thr Phe

1

5

10

15

Pro Val Ala Met Phe Trp Val Ser Asn Gln Ala Glu Trp Phe Glu Asp

35 40 45
Glu Ile Glu Glu Phe Lys Glu Arg Leu Arg Lys Arg Arg Glu Glu Lys
50 55 60
Leu Leu Arg Asp Ala Gln Gln Asn Ser
65 70

Sequence No.: 7

Sequence length: 1174

Sequence type: Amino acid

Topology: Linear

Sequence kind: Protein

Hypothetical: No

Original source:

Organism species: *Homo sapiens*

Cell kind: Histiocyte lymphoma

Cell line: U937

Clone name: HP10269

Sequence description

Met Arg Pro Phe Phe Leu Leu Cys Phe Ala Leu Pro Gly Leu Leu His
1 5 10 15
Ala Gln Gln Ala Cys Ser Arg Gly Ala Cys Tyr Pro Pro Val Gly Asp
20 25 30
Leu Leu Val Gly Arg Thr Arg Phe Leu Arg Ala Ser Ser Thr Cys Gly
35 40 45
Leu Thr Lys Pro Glu Thr Tyr Cys Thr Gln Tyr Gly Glu Trp Gln Met
50 55 60
Lys Cys Cys Lys Cys Asp Ser Arg Gln Pro His Asn Tyr Tyr Ser His
70 75 80

73

85	90	95
Ser Gln Asn Asp Val Asn Pro Val Ser Leu Gln Leu Asp Leu Asp Arg		
100	105	110
Arg Phe Gln Leu Gln Glu Val Met Met Glu Phe Gln Gly Pro Met Pro		
115	120	125
Ala Gly Met Leu Ile Glu Arg Ser Ser Asp Phe Gly Lys Thr Trp Arg		
130	135	140
Val Tyr Gln Tyr Leu Ala Ala Asp Cys Thr Ser Thr Phe Pro Arg Val		
145	150	155
Arg Gln Gly Arg Pro Gln Ser Trp Gln Asp Val Arg Cys Gln Ser Leu		
165	170	175
Pro Gln Arg Pro Asn Ala Arg Leu Asn Gly Gly Lys Val Gln Leu Asn		
180	185	190
Leu Met Asp Leu Val Ser Gly Ile Pro Ala Thr Gln Ser Gln Lys Ile		
195	200	205
Gln Glu Val Gly Glu Ile Thr Asn Leu Arg Val Asn Phe Thr Arg Leu		
210	215	220
Ala Pro Val Pro Gln Arg Gly Tyr His Pro Pro Ser Ala Tyr Tyr Ala		
225	230	235
Val Ser Gln Leu Arg Leu Gln Gly Ser Cys Phe Cys His Gly His Ala		
245	250	255
Asp Arg Cys Ala Pro Lys Pro Gly Ala Ser Ala Gly Pro Ser Thr Ala		
260	265	270
Val Gln Val His Asp Val Cys Val Cys Gln His Asn Thr Ala Gly Pro		
275	280	285
Asn Cys Glu Arg Cys Ala Pro Phe Tyr Asn Asn Arg Pro Trp Arg Pro		
290	295	300

74

His Ser Glu Thr Cys His Phe Asp Pro Ala Val Phe Ala Ala Ser Gln

325

330

335

Gly Ala Tyr Gly Gly Val Cys Asp Asn Cys Arg Asp His Thr Glu Gly

340

345

350

Lys Asn Cys Glu Arg Cys Gln Leu His Tyr Phe Arg Asn Arg Arg Pro

355

360

365

Gly Ala Ser Ile Gln Glu Thr Cys Ile Ser Cys Glu Cys Asp Pro Asp

370

375

380

Gly Ala Val Pro Gly Ala Pro Cys Asp Pro Val Thr Gly Gln Cys Val

385

390

395

400

Cys Lys Glu His Val Gln Gly Glu Arg Cys Asp Leu Cys Lys Pro Gly

405

410

415

Phe Thr Gly Leu Thr Tyr Ala Asn Pro Gln Gly Cys His Arg Cys Asp

420

425

430

Cys Asn Ile Leu Gly Ser Arg Arg Asp Met Pro Cys Asp Glu Glu Ser

435

440

445

Gly Arg Cys Leu Cys Leu Pro Asn Val Val Gly Pro Lys Cys Asp Gln

450

455

460

Cys Ala Pro Tyr His Trp Lys Leu Ala Ser Gly Gln Gly Cys Glu Pro

465

470

475

480

Cys Ala Cys Asp Pro His Asn Ser Leu Ser Pro Gln Cys Asn Gln Phe

485

490

495

Thr Gly Gln Cys Pro Cys Arg Glu Gly Phe Gly Gly Leu Met Cys Ser

500

505

510

Ala Ala Ala Ile Arg Gln Cys Pro Asp Arg Thr Tyr Gly Asp Val Ala

515

520

525

Thr Gly Cys Arg Ala Cys Asp Cys Asp Phe Arg Gly Thr Glu Gly Pro

75

545	550	555	560
Gly Pro Arg Cys Asp Gln Cys Gln Arg Gly Tyr Cys Asn Arg Tyr Pro			
565	570	575	
Val Cys Val Ala Cys His Pro Cys Phe Gln Thr Tyr Asp Ala Asp Leu			
580	585	590	
Arg Glu Gln Ala Leu Arg Phe Gly Arg Leu Arg Asn Ala Thr Ala Ser			
595	600	605	
Leu Trp Ser Gly Pro Gly Leu Glu Asp Arg Gly Leu Ala Ser Arg Ile			
610	615	620	
Leu Asp Ala Lys Ser Lys Ile Glu Gln Ile Arg Ala Val Leu Ser Ser			
625	630	635	640
Pro Ala Val Thr Glu Gln Glu Val Ala Gln Val Ala Ser Ala Ile Leu			
645	650	655	
Ser Leu Arg Arg Thr Leu Gln Gly Leu Gln Leu Asp Leu Pro Leu Glu			
660	665	670	
Glu Glu Thr Leu Ser Leu Pro Arg Asp Leu Glu Ser Leu Asp Arg Ser			
675	680	685	
Phe Asn Gly Leu Leu Thr Met Tyr Gln Arg Lys Arg Glu Gln Phe Glu			
690	695	700	
Lys Ile Ser Ser Ala Asp Pro Ser Gly Ala Phe Arg Met Leu Ser Thr			
705	710	715	720
Ala Tyr Glu Gln Ser Ala Gln Ala Ala Gln Gln Val Ser Asp Ser Ser			
725	730	735	
Arg Leu Leu Asp Gln Leu Arg Asp Ser Arg Arg Glu Ala Glu Arg Leu			
740	745	750	
Val Arg Gln Ala Gly Gly Gly Gly Thr Gly Ser Pro Lys Leu Val			
755	760	765	

Asn Lys Leu Cys Gly Asn Ser Arg Gln Met Ala Cys Thr Pro Ile Ser
 785 790 795 800
 Cys Pro Gly Glu Leu Cys Pro Gln Asp Asn Gly Thr Ala Cys Gly Ser
 805 810 815
 Arg Cys Arg Gly Val Leu Pro Arg Ala Gly Gly Ala Phe Leu Met Ala
 820 825 830
 Gly Gln Val Ala Glu Gln Leu Arg Gly Phe Asn Ala Gln Leu Gln Arg
 835 840 845
 Thr Arg Gln Met Ile Arg Ala Ala Glu Glu Ser Ala Ser Gln Ile Gln
 850 855 860
 Ser Ser Ala Gln Arg Leu Glu Thr Gln Val Ser Ala Ser Arg Ser Gln
 865 870 875 880
 Met Glu Glu Asp Val Arg Arg Thr Arg Leu Leu Ile Gln Gln Val Arg
 885 890 895
 Asp Phe Leu Thr Asp Pro Asp Thr Asp Ala Ala Thr Ile Gln Glu Val
 900 905 910
 Ser Glu Ala Val Leu Ala Leu Trp Leu Pro Thr Asp Ser Ala Thr Val
 915 920 925
 Leu Gln Lys Met Asn Glu Ile Gln Ala Ile Ala Ala Arg Leu Pro Asn
 930 935 940
 Val Asp Leu Val Leu Ser Gln Thr Lys Gln Asp Ile Ala Arg Ala Arg
 945 950 955 960
 Arg Leu Gln Ala Glu Ala Glu Glu Ala Arg Ser Arg Ala His Ala Val
 965 970 975
 Glu Gly Gln Val Glu Asp Val Val Gly Asn Leu Arg Gln Gly Thr Val
 980 985 990
 Ala Leu Gln Glu Ala Gln Asp Thr Met Gln Gly Thr Ser Arg Ser Leu

1010	1015	1020	
Ala Glu Lys Leu Val Thr Ser Met Thr Lys Gln Leu Gly Asp Phe Trp			
1025	1030	1035	1040
Thr Arg Met Glu Glu Leu Arg His Gln Ala Arg Gln Gln Gly Ala Glu			
1045	1050	1055	
Ala Val Gln Ala Gln Gln Leu Ala Glu Gly Ala Ser Glu Gln Ala Leu			
1060	1065	1070	
Ser Ala Gln Glu Gly Phe Glu Arg Ile Lys Gln Lys Tyr Ala Glu Leu			
1075	1080	1085	
Lys Asp Arg Leu Gly Gln Ser Ser Met Leu Gly Glu Gln Gly Ala Arg			
1090	1095	1100	
Ile Gln Ser Val Lys Thr Glu Ala Glu Glu Leu Phe Gly Glu Thr Met			
1105	1110	1115	1120
Glu Met Met Asp Arg Met Lys Asp Met Glu Leu Glu Leu Leu Arg Gly			
1125	1130	1135	
Ser Gln Ala Ile Met Leu Arg Ser Ala Asp Leu Thr Gly Leu Glu Lys			
1140	1145	1150	
Arg Val Glu Gln Ile Arg Asp His Ile Asn Gly Arg Val Leu Tyr Tyr			
1155	1160	1165	
Ala Thr Cys Lys			
1170			

Sequence No.: 8

Sequence length: 122

Sequence type: Amino acid

Topology: Linear

Sequence kind: Protein

Original source

Organism species: *Homo sapiens*

Cell kind: Stomach cancer

Clone name: HP10298

Sequence description

Met Gly Leu Leu Leu Leu Val Pro Leu Leu Leu Leu Pro Gly Ser Tyr

1 5 10 15

Gly Leu Pro Phe Tyr Asn Gly Phe Tyr Tyr Ser Asn Ser Ala Asn Asp

20 25 30

Gln Asn Leu Gly Asn Gly His Gly Lys Asp Leu Leu Asn Gly Val Lys

35 40 45

Leu Val Val Glu Thr Pro Glu Glu Thr Leu Phe Thr Arg Ile Leu Thr

50 55 60

Val Gly Pro Gln Ser Leu Gly Ser Glu Ala Leu Ala Ser Pro Thr Arg

65 70 75 80

Arg Ala Ala Cys Thr Val Phe Thr Ala Thr Ala Ser Thr Arg Thr Trp

85 90 95

Gly Pro Pro Leu Pro His Ser Leu Thr Gly Cys Val Phe Ile Glu Trp

100 105 110

Phe Val Phe Pro Cys Gly Leu Glu Pro Phe

115 120

Sequence No.: 9

Sequence length: 175

Sequence type: Amino acid

Topology: Linear

Sequence kind: Protein

Hypothetical: No

Cell kind: Stomach cancer

Clone name: HP10368

Sequence description

Met Glu Lys Ile Pro Val Ser Ala Phe Leu Leu Leu Val Ala Leu Ser
 1 5 10 15
 Tyr Thr Leu Ala Arg Asp Thr Thr Val Lys Pro Gly Ala Lys Lys Asp
 20 25 30
 Thr Lys Asp Ser Arg Pro Lys Leu Pro Gln Thr Leu Ser Arg Gly Trp
 35 40 45
 Gly Asp Gln Leu Ile Trp Thr Gln Thr Tyr Glu Glu Ala Leu Tyr Lys
 50 55 60
 Ser Lys Thr Ser Asn Lys Pro Leu Met Ile Ile His His Leu Asp Glu
 65 70 75 80
 Cys Pro His Ser Gln Ala Leu Lys Lys Val Phe Ala Glu Asn Lys Glu
 85 90 95
 Ile Gln Lys Leu Ala Glu Gln Phe Val Leu Leu Asn Leu Val Tyr Glu
 100 105 110
 Thr Thr Asp Lys His Leu Ser Pro Asp Gly Gln Tyr Val Pro Arg Ile
 115 120 125
 Met Phe Val Asp Pro Ser Leu Thr Val Arg Ala Asp Ile Thr Gly Arg
 130 135 140
 Tyr Ser Asn Arg Leu Tyr Ala Tyr Glu Pro Ala Asp Thr Ala Leu Leu
 145 150 155 160
 Leu Asp Asn Met Lys Lys Ala Leu Lys Leu Leu Lys Thr Glu Leu
 165 170 175

Sequence No.: 10

Sequence type: Nucleic acid

Strandedness: Double

Topology: Linear

Sequence kind: cDNA to mRNA

Original source:

Organism species: *Homo sapiens*

Cell kind: Fibrosarcoma

Cell line: HT-1080

Clone name: HP00658

Sequence description

ATGAAGGTCT CCGCGGCAGC CCTCGCTGTC ATCCTCATTG CTA CTGCGCTCCT	60
GCATCTGCCT CCCCATATTC CTCGCACACC ACACCCTGCT GCTTTGCCTA CATTGCCCGC	120
CCACTGCCCC GTGCCCACAT CAAGGAGTAT TTCTACACCA GTGGCAAGTG CTCCAACCCA	180
GCAGTCGTCC ACAGGTCAAG GATGCCAAAG AGAGAGGGAC AGCAAGTCTG GCAGGATTTT	240
CTGTATGACT CCCGGCTGAA CAAGGGCAAG CTTTGTACCC CGAAAGAACC GCCAAGTGTG	300
TGCCAACCCA GAGAAGAAAT GGGTTCGGGA GTACATCAAC TCTTTGGAGA TGAGCTAGGA	360
TGGAGAGTCC TTGAACCTGA ACTTACACAA ATTTGCCTGT TTCTGCTTGC TCTTGTCTTA	420
GCTTGGGAGG CTTCCCTCA CTATCCTACC CCACCCGCTC CT	462

Sequence No.: 11

Sequence length: 945

Sequence type: Nucleic acid

Strandedness: Double

Topology: Linear

Sequence kind: cDNA to mRNA

Original source:

Organism species: *Homo sapiens*

Clone name: HP00714

Sequence description

ATGGACCTGC GACAGTTTCT TATGTGCCTG TCCCTGTGCA CAGCCTTTGC CTTGAGCAAA	60
CCCACAGAAA AGAAGGACCG TGTACATCAT GAGCCTCAGC TCAGTGACAA GGTTCACAAT	120
GATGCTCAGA GTTTTGATTA TGACCATGAT GCCTTCTTGG GTGCTGAAGA AGCAAAGACC	180
TTTGATCAGC TGACACCAGA AGAGAGCAAG GAAAGGCTTG GAAAGATTGT AAGTAAAATA	240
GATGCTCAGA AGGACGGGTT TGTCACTGTG GATGAGCTCA AAGACTGGAT TAAATTTGCA	300
CAAAAGCGCT GGATTTACGA GGATGTAGAG CGACAGTGGG AGGGGCATGA CCTCAATGAG	360
GACGGCCTCG TTTCTGGGA GGAGTATAAA AATGCCACCT ACGGCTACGT TTTAGATGAT	420
CCAGATCCTG ATGATGGATT TAACTATAAA CAGATGATGG TTAGAGATGA GCGGAGGTTT	480
AAAATGGCAG ACAAGGATGG AGACCTCATT GCCACCAAGG AGGAGTTCAC AGCTTTCTCTG	540
CACCCTGAGG AGTATGACTA CATGAAAGAT ATAGTAGTAC AGGAAACAAT GGAAGATATA	600
GATAAGAATG CTGATGGTTT CATTGATCTA GAAGAGTATA TTGGTGACAT GTACAGCCAT	660
GATGGGAATA CTGATGAGCC AGAATGGGTA AAGACAGAGC GAGAGCAGTT TGTTGAGTTT	720
CGGGATAAGA ACCGTGATGG GAAGATGGAC AAGGAAGAGA CCAAAGACTG GATCCTTCCC	780
TCAGACTATG ATCATGCAGA GGCAGAAGCC AGGCACCTGG TCTATGAATC AGACCAAAAC	840
AAGGATGGCA AGCTTACCAA GGAGGAGATC GTTGACAAGT ATGACTTATT TGTGGCAGC	900
CAGGCCACAG ATTTTGGGGA GGCCTTAGTA CGGCATGATG AGTTC	945

Sequence No.: 12

Sequence length: 474

Sequence type: Nucleic acid

Strandedness: Double

Topology: Linear

Sequence kind: cDNA to mRNA

Original source:

Organism species: *Homo sapiens*

Clone name: HP00871

Sequence description

ATGGCTTCCA GAAGCATGCG GCTGCTCCTA TTGCTGAGCT GCCTGGCCAA AACAGGAGTC 60
CTGGGTGATA TCATCATGAG ACCCAGCTGT GCTCCTGGAT GGTTTTACCA CAAGTCCAAT 120
TGCTATGGTT ACTTCAGGAA GCTGAGGAAC TGGTCTGATG CCGAGCTCGA GTGTCACTCT 180
TACGGAAACG GAGCCACCT GGCATCTATC CTGAGTTTAA AGGAAGCCAG CACCATAGCA 240
GAGTACATAA GTGGCTATCA GAGAAGCCAG CCGATATGGA TTGGCCTGCA CGACCCACAG 300
AAGAGCCAGG ACTGCGAGTG GATTGATGGG GCCATGTATC TGTACAGATC CTGGTCTGGC 360
AAGTCCATGG GTGGGAACAA GCACTGTGCT GAGATGAGCT CCAATAACAA CTTTTTAACT 420
TGGAGCAGCA ACGAATGCAA CAAGCGCCAA CACTTCCTGT GCAAGTACCG ACCA 474

Sequence No.: 13

Sequence length: 1128

Sequence type: Nucleic acid

Strandedness: Double

Topology: Linear

Sequence kind: cDNA to mRNA

Original source:

Organism species: *Homo sapiens*

Cell kind: Liver

Clone name: HP01134

Sequence description

ATGGTTTGA AAGTAGCTGT ATTCCTCAGT GTGGCCCTGG GCATTGGTGC CGTTCCTATA 60
GATGATCCTG AAGATGGAGG CAAGCACTGG GTGGTGATCG TGGCAGGTTT AAATGGCTGG 120
TATAATTATA GGCACCAGGC AGACGCGTGC CATGCCCTACC AGATCATTCA CCGCAATGGG 180
ATTCCTGACG AACAGATCGT TGTGATGATG TACGATGACA TTGCTTACTC TGAAGACAAT 240
CCCACTCCAG GAATTGTGAT CAACAGGCCC AATGGCACAG ATGTCTATCA GGGAGTCCCC 300
AAGGACTACA CTGGAGAGGA TGTTACCCCA CAAAATTTC TTGCTGTGTT GAGAGGCGAT 360

GTGTTTCAAT ACTTCACTGA CCATGGATCT ACTGGATG

CTTCATGTAA AGGACCTGAA TGAGACCATC CATTACATGT ACAAACACAA AATGTACCGA 540
AAGATGGTGT TCTACATTGA AGCCTGTGAG TCTGGGTCCA TGATGAACCA CCTGCCGGAT 600
AACATCAATG TTTATGCAAC TACTGCTGCC AACCCAGAG AGTCGTCCTA GGCCTGTTAC 660
TATGATGAGA AGAGGTCCAC GTACCTGGGG GACTGGTACA GCGTCAACTG GATGGAAGAC 720
TCGGACGTGG AAGATCTGAC TAAAGAGACC CTGCACAAGC AGTACCACCT GGTAAAATCG 780
CACACCAACA CCAGCCACGT CATGCAGTAT GGAAACAAAA CAATCTCCAC CATGAAAGTG 840
ATGCAGTTTC AGGGTATGAA ACGCAAAGCC AGTTCTCCCG TCCCCTACC TCCAGTCACA 900
CACCTTGACC TCACCCCCAG CCCTGATGTG CCTCTACCA TCATGAAAAG GAAACTGATG 960
AACACCAATG ATCTGGAGGA GTCCAGGCAG CTCACGGAGG AGATCCAGCG GCATCTGGAT 1020
TACGAGTATG CGTTGAGACA TTTGTACGTG CTGGTCAACC TTTGTGAGAA GCCGTATCCG 1080
CTTCACAGGA TAAATTGTC CATGGACCAC GTGTGCCTTG GTCACTAC 1128

Sequence No.: 14

Sequence length: 519

Sequence type: Nucleic acid

Strandedness: Double

Topology: Linear

Sequence kind: cDNA to mRNA

Original source:

Organism species: *Homo sapiens*

Cell kind: Epidermoid carcinoma

Cell line: KB

Clone name: HP10029

Sequence description

ATGGCGGCGC CCAGCGGAGG GTGGAACGGC GTCCGCGCGA GCTTGTGGGC CGCGCTGCTC 60
CTAGGGGCGG TGGCGCTGAG GCCGGCGGAG GCGGTGTCCG AGCCCACGAC CGTGGCGTTT 120
GACGTGCGGC CCGGCGGCGT CGTGCAATTCC TTCTCCATA ACGTGGGCCC GGGGGACAAA 180

ATGGGGAACG GCGGAAGACC GCGGACCG

TCCTATCTGT ACTTCACACA GTTCAAGGCA GAGGTGCGGG GCGCTGAGAT TGAGTACGCC 360
ATGGCCTACT CTAAAGCCGC ATTTGAAAGG GAAAGTGATG TCCCTCTGAA AACTGAGGAA 420
TTTGAAGTGA CCAAAACAGC AGTGGCTCAC AGGCCCGGGG CATTCAAAGC TGAGCTGTCC 480
AAGCTGGTGA TTGTGGCCAA GGCATCGCGC ACTGAGCTG 519

Sequence No.: 15

Sequence length: 219

Sequence type: Nucleic acid

Strandedness: Double

Topology: Linear

Sequence kind: cDNA to mRNA

Original source:

Organism species: *Homo sapiens*

Cell kind: Epidermoid carcinoma

Cell line: KB

Clone name: HP10189

Sequence description

ATGGGGGTGA AGCTGGAGAT ATTTCCGATG ATAATCTACC TCACTTTCCC TGTGGCTATG 60
TTCTGGGTTT CCAATCAGGC CGAGTGGTTT GAGGACGATG TCATACAGCG CAAGAGGGAG 120
CTGTGGCCAC CTGAGAAGCT TCAAGAGATA GAGGAATTCA AAGAGAGGTT ACGGAAGCGG 180
CGGGAGGAGA AGCTCCTTCG CGACGCCCAG CAGAACTCC 219

Sequence No.: 16

Sequence length: 3516

Sequence type: Nucleic acid

Strandedness: Double

Topology: Linear

Organism species: *Homo sapiens*

Cell kind: Lymphoma

Cell line: U937

Clone name: HP10269

Sequence description

```
ATGAGACCAT TCTTCTCTT GTGTTTGGCC CTGCCTGGCC TCCTGCATGC CCAACAAGCC      60
TCCTCCCGTG GGGCCTGCTA TCCACCTGTT GGGGACCTGC TTGTGGGAG GACCCGGTTT      120
CTCCGAGCTT CATCTACCTG TGGACTGACC AAGCCTGAGA CCTACTGCAC CCAGTATGGC      180
GAGTGGCAGA TGAAATGCTG CAAGTGTGAC TCCAGGCAGC CTCACAACCTA CTACAGTCAC      240
CGAGTAGAGA ATGTGGCTTC ATCCTCCGGC CCCATGCGCT GGTGGCAGTC CCAGAATGAT      300
GTGAACCTG TCTCTCTGCA GCTGGACCTG GACAGGAGAT TCCAGCTTCA AGAAGTCATG      360
ATGGAGTTCC AGGGGCCCAT GCCTGCCGGC ATGCTGATTG AGCGCTCCTC AGACTTCGGT      420
AAGACCTGGC GAGTGTACCA GTACCTGGCT GCCGACTGCA CCTCCACCTT CCCTCGGGTC      480
CGCCAGGGTC GGCCTCAGAG CTGGCAGGAT GTTCGGTGCC AGTCCCTGCC TCAGAGGCCT      540
AATGCACGCC TAAATGGGGG GAAGGTCCAA CTTAACCTTA TGGATTTAGT GTCTGGGATT      600
CCAGCAACTC AAAGTCAAAA AATTCAAGAG GTGGGGGAGA TCACAAACTT GAGAGTCAAT      660
TTCACCAGGC TGGCCCCTGT GCCCCAAAGG GGCTACCACC CTCCCAGCGC CTAATATGCT      720
GTGTCCCAGC TCCGTCTGCA GGGGAGCTGC TTCTGTCACG GCCATGCTGA TCGCTGCGCA      780
CCCAAGCCTG GGGCCTCTGC AGGCCCCTCC ACCGCTGTGC AGGTCCACGA TGTCTGTGTC      840
TGCCAGCACA ACACTGCCGG CCCAAATTGT GAGCGCTGTG CACCCTTCTA CAACAACCGG      900
CCCTGGAGAC CGGCGGAGGG CCAGGACGCC CATGAATGCC AAAGGTGCGA CTGCAATGGG      960
CACTCAGAGA CATGTCACCT TGACCCCGCT GTGTTTGCCG CCAGCCAGGG GGCATATGGA     1020
GGTGTGTGTG ACAATTGCCG GGACCACACC GAAGGCAAGA ACTGTGAGCG GTGTCAGCTG     1080
CACTATTTCC GGAACCGGGC CCCGGGAGCT TCCATTGAGG AGACCTGCAT CTCCTGCGAG     1140
TGTGATCCGG ATGGGGCAGT GCCAGGGGCT CCCTGTGACC CAGTGACCGG GCAGTGTGTG     1200
TGCAAGGAGC ATGTGCAGGG AGAGCGCTGT GACCTATGCA AGCCGGGCTT CACTGGACTC     1260
ACCTACGCCA ACCCGCAGGG CTGCCACCGC TGTGACTGCA ACATCCTGGG GTCCCGGAGG     1320
```

AAATGTGACC AGTGTGCTG TACCACTG GAGGCTG

TGTGCCTGCG ACCCGCACAA CTCCCTCAGC CCACAGTGCA ACCAGTTCAC AGGGCAGTGC 1500
CCCTGTGCGG AAGGCTTTGG TGGCCTGATG TGCAGCGCTG CAGCCATCCG CCAGTGTCCA 1560
GACCGGACCT ATGGAGACGT GGCCACAGGA TGCCGAGCCT GTGACTGTGA TTTCCGGGGA 1620
ACAGAGGGCC CGGGCTGCGA CAAGGCATCA GGCCGCTGCC TCTGCCGCCC TGGCTTGACC 1680
GGGCCCCGCT GTGACCAGTG CCAGCGAGGC TACTGCAATC GCTACCCGGT GTGCGTGGCC 1740
TGCCACCCTT GCTTCCAGAC CTATGATGCG GACCTCCGGG AGCAGGCCCT GCGCTTTGGT 1800
ACACTCCGCA ATGCCACCGC CAGCCTGTGG TCAGGGCCTG GGCTGGAGGA CCGTGGCCTG 1860
GCCTCCCGGA TCCTAGATGC AAAGAGTAAG ATTGAGCAGA TCCGAGCAGT TCTCAGCAGC 1920
CCCGCAGTCA CAGAGCAGGA GGTGGCTCAG GTGGCCAGTG CCATCCTCTC CCTCAGGCGA 1980
ACTCTCCAGG GCCTGCAGCT GGATCTGCCC CTGGAGGAGG AGACGTTGTC CCTTCCGAGA 2040
GACCTGGAGA GTCTTGACAG AAGCTTCAAT GGTCTCCTTA CTATGTATCA GAGGAAGAGG 2100
GAGCAGTTTG AAAAAATAAG CAGTGCTGAT CCTTCAGGAG CCTTCCGGAT GCTGAGCACA 2160
GCCTACGAGC AGTCAGCCCA GGCTGCTCAG CAGGTCTCCG ACAGCTCGCG CCTTTTGGAC 2220
CAGCTCAGGG ACAGCCGGAG AGAGGCAGAG AGGCTGGTGC GGCAGGCGGG AGGAGGAGGA 2280
GGCACCGGCA GCCCCAAGCT TGTGGCCCTG AGGCTGGAGA TGTCTTGGTT GCCTGACCTG 2340
ACACCCACCT TCAACAAGCT CTGTGGCAAC TCCAGGCAGA TGGCTTGAC CCAATATCA 2400
TGCCCTGGTG AGCTATGTCC CCAAGACAAT GGCACAGCCT GTGGCTCCCG CTGCAGGGGT 2460
GTCCTTCCCA GGGCCGGTGG GGCCTTCTTG ATGGCGGGGC AGGTGGCTGA GCAGCTGCGG 2520
GGCTTCAATG CCCAGCTCCA GCGGACCAGG CAGATGATTA GGGCAGCCGA GGAATCTGCC 2580
TCACAGATTC AATCCAGTGC CCAGCGCTTG GAGACCCAGG TGAGCGCCAG CCGCTCCAG 2640
ATGGAGGAAG ATGTCAGACG CACACGGCTC CTAATCCAGC AGGTCCGGGA CTTCTAACA 2700
GACCCCGACA CTGATGCAGC CACTATCCAG GAGGTCAGCG AGGCCGTGCT GGCCCTGTGG 2760
CTGCCCACAG ACTCAGCTAC TGTTCGACAG AAGATGAATG AGATCCAGGC CATTGCAGCC 2820
AGGCTCCCCA ACGTGGACTT GGTGCTGTCC CAGACCAAGC AGGACATTGC GCGTGCCCGC 2880
CGGTTGCAGG CTGAGGCTGA GGAAGCCAGG AGCCGAGCCC ATGCAGTGGA GGGCCAGGTG 2940
GAAGATGTGG TTGGGAACCT GCGGCAGGGG ACAGTGGCAC TGCAGGAAGC TCAGGACACC 3000
ATGCAAGGCA CCAGCCGCTC CCTTCGGCTT ATCCAGGACA GGTTGCTGA GGTTCAGCAG 3060

AAACGGATGG AGGAAGTGGG GAGCCAAAGC

87

CAGCAGCTTG CGGAAGGTGC CAGCGAGCAG GCATTGAGTG CCCAAGAGGG ATTTGAGAGA 3240
ATAAAACAAA AGTATGCTGA GTTGAAGGAC CGGTTGGGTC AGAGTTCCAT GCTGGGTGAG 3300
CAGGGTGCCC GGATCCAGAG TGTGAAGACA GAGGCAGAGG AGCTGTTTGG GGAGACCATG 3360
GAGATGATGG ACAGGATGAA AGACATGGAG TTGGAGCTGC TCGGGGGCAG CCAGGCCATC 3420
ATGCTGCGCT CAGCGGACCT GACAGGACTG GAGAAGCGTG TGGAGCAGAT CCGTGACCAC 3480
ATCAATGGGC GCGTGCTCTA CTATGCCACC TGCAAG 3516

Sequence No.: 17

Sequence length: 366

Sequence type: Nucleic acid

Strandedness: Double

Topology: Linear

Sequence kind: cDNA to mRNA

Original source:

Organism species: *Homo sapiens*

Cell kind: Stomach cancer

Clone name: HP10298

Sequence description

ATGGGCCTGT TGCTCCTGGT CCCATTGCTC CTGCTGCCCG GCTCCTACGG ACTGCCCTTC 60
TACAACGGCT TCTACTACTC CAACAGCGCC AACGACCAGA ACCTAGGCAA CGGTCATGGC 120
AAAGACCTCC TTAATGGAGT GAAGCTGGTG GTGGAGACAC CCGAGGAGAC CCTGTTCAAC 180
CGCATCCTAA CTGTGGGCCC CCAGAGCCTG GGGTCCGAAG CTTTGGCTTC CCCGACCCGC 240
AGAGCCGCTT GTACGGTGTT TACTGCTACC GCCAGCACTA GGACCTGGGG CCCTCCCCTG 300
CCGCATTCCC TCACTGGCTG TGTATTTATT GAGTGGTTCG TTTTCCCTTG TGGGTTGGAG 360
CCATTT 366

Sequence No.: 18

Sequence type: Nucleic acid

Strandedness: Double

Topology: Linear

Sequence kind: cDNA to mRNA

Original source:

Organism species: *Homo sapiens*

Cell kind: Stomach cancer

Clone name: HP10368

Sequence description

ATGGAGAAAA TTCCAGTGTC AGCATTCTTG CTCCTTGTTG CCCTCTCCTA CACTCTGGCC	60
AGAGATACCA CAGTCAAACC TGGAGCCAAA AAGGACACAA AGGACTCTCG ACCCAAACCTG	120
CCCCAGACCC TCTCCAGAGG TTGGGGTGAC CAACTCATCT GGA CTCAGAC ATATGAAGAA	180
GCTCTATATA AATCCAAGAC AAGCAACAAA CCCTTGATGA TTATTCATCA CTTGGATGAG	240
TGCCCCACACA GTCAAGCTTT AAAGAAAGTG TTTGCTGAAA ATAAAGAAAT CCAGAAATTG	300
GCAGAGCAGT TTGTCTCCT CAATCTGGTT TATGAAACAA CTGACAAACA CCTTTCTCCT	360
GATGGCCAGT ATGTCCCCAG GATTATGTTT GTTGACCCAT CTCTGACAGT TAGAGCCGAT	420
ATCACTGGAA GATATTCAAA CCGTCTCTAT GCTTACGAAC CTGCAGATAC AGCTCTGTTG	480
CTTGACAACA TGAAGAAAGC TCTCAAGTTG CTGAAGACTG AATTG	525

Sequence No.: 19

Sequence length: 1296

Sequence type: Nucleic acid

Strandedness: Double

Topology: Linear

Sequence kind: cDNA to mRNA

Original source:

Organism species: *Homo sapiens*

Cell kind: Fibrosarcoma

Sequence characteristics:

Code representing characteristics: CDS

Existence site: 56.. 520

Characterization method: E

Sequence description

CCTGCAGAGG ATCAAGACAG CACGTGGACC TCGCACAGCC TCTCCCACAG GTACC ATG	58
Met	
1	
AAG GTC TCC GCG GCA GCC CTC GCT GTC ATC CTC ATT GCT ACT GCC CTC	106
Lys Val Ser Ala Ala Ala Leu Ala Val Ile Leu Ile Ala Thr Ala Leu	
5 10 15	
TGC GCT CCT GCA TCT GCC TCC CCA TAT TCC TCG GAC ACC ACA CCC TGC	154
Cys Ala Pro Ala Ser Ala Ser Pro Tyr Ser Ser Asp Thr Thr Pro Cys	
20 25 30	
TGC TTT GCC TAC ATT GCC CGC CCA CTG CCC CGT GCC CAC ATC AAG GAG	202
Cys Phe Ala Tyr Ile Ala Arg Pro Leu Pro Arg Ala His Ile Lys Glu	
35 40 45	
TAT TTC TAC ACC AGT GGC AAG TGC TCC AAC CCA GCA GTC GTC CAC AGG	250
Tyr Phe Tyr Thr Ser Gly Lys Cys Ser Asn Pro Ala Val Val His Arg	
50 55 60 65	
TCA AGG ATG CCA AAG AGA GAG GGA CAG CAA GTC TGG CAG GAT TTC CTG	298
Ser Arg Met Pro Lys Arg Glu Gly Gln Gln Val Trp Gln Asp Phe Leu	
70 75 80	
TAT GAC TCC CGG CTG AAC AAG GGC AAG CTT TGT CAC CCG AAA GAA CCG	346
Tyr Asp Ser Arg Leu Asn Lys Gly Lys Leu Cys His Pro Lys Glu Pro	
85 90 95	
CCA AGT GTG TGC CAA CCC AGA GAA GAA ATG GGT TCG GGA GTA CAT CAA	394

90

CTC TTT GGA GAT GAG CTA GGA TGG AGA GTC CTT GAA CCT GAA CTT ACA 442
Leu Phe Gly Asp Glu Leu Gly Trp Arg Val Leu Glu Pro Glu Leu Thr
115 120 125
CAA ATT TGC CTG TTT CTG CTT GCT CTT GTC CTA GCT TGG GAG GCT TCC 490
Gln Ile Cys Leu Phe Leu Leu Ala Leu Val Leu Ala Trp Glu Ala Ser
130 135 140 145
CCT CAG TAT CCT ACC CCA CCC GCT CCT TGAAGGGCCC AGA 530
Pro His Tyr Pro Thr Pro Pro Ala Pro
150
TTCTACCACA CAGCAGCAGT TACAAAAACC TTCCCAGGC TGGACGTGGT GGCTCACGCC 590
TGTAATCCCA GCACTTTGGG AGGCCAAGGT GGTGGATCA CTTGAGGTCA GGAGTTCGAG 650
ACCAGCCTGG CCAACATGAT GAAACCCCAT CTCTACTAAA AATACAAAAA ATTAGCCGGG 710
CGTGGTAGCG GGCGCCTGTA GTCCCAGCTA CTCGGGAGGC TGAGGCAGGA GAATGGCGTG 770
AACCCGGGAG GCGGAGCTTG CAGTGAGCCG AGATCGCGCC ACTGCACTCC AGCCTGGGCG 830
ACAGAGCGAG ACTCCGTCTC AAAAAAAAAA AAAAAAAAAA AAATACAAAA ATTAGCCGGG 890
CGTGGTGGCC CACGCCTGTA ATCCCAGCTA CTCGGGAGGC TAAGGCAGGA AAATTGTTTG 950
AACCCAGGAG GTGGAGGCTG CAGTGAGCTG AGATTGTGCC ACTTCACTCC AGCCTGGGTG 1010
ACAAAGTGAG ACTCCGTAC AACAACAACA AAAAAAGCT TCCCCAACTA AAGCCTAGAA 1070
GAGCTTCTGA GGCGCTGCTT TGTCAAAAGG AAGTCTCTAG GTTCTGAGCT CTGGCTTTGC 1130
CTTGGCTTTG CCAGGGCTCT GTGACCAGGA AGGAAGTCAG CATGCCTCTA GAGGCAAGGA 1190
GGGGAGGAAC GCTGCACTCT TAAGCTTCG CCGTCTCAAC CCCTCACAGG AGCTTACTGG 1250
CAAACATGAA AAATCGGCTT ACCATTAAAG TTCTCAATGC AACCAT 1296

Sequence No.: 20

Sequence length: 3311

Sequence type: Nucleic acid

Strandedness: Double

SEQUENCE KIND: cDNA

Original source:

Organism species: *Homo sapiens*
Cell kind: Epidermoid carcinoma
Cell line: KB
Clone name: HP00714

Sequence characteristics:

Code representing characteristics: CDS

Existence site: 57.. 1004

Characterization method: E

Sequence description

GAGCGGCGGC CACGGCATCC TGTGCTGTGG GGGCTACGAG GAAAGATCTA ATTATC ATG 59
Met
1
GAC CTG CGA CAG TTT CTT ATG TGC CTG TCC CTG TGC ACA GCC TTT GCC 107
Asp Leu Arg Gln Phe Leu Met Cys Leu Ser Leu Cys Thr Ala Phe Ala
5 10 15
TTG AGC AAA CCC ACA GAA AAG AAG GAC CGT GTA CAT CAT GAG CCT CAG 155
Leu Ser Lys Pro Thr Glu Lys Lys Asp Arg Val His His Glu Pro Gln
20 25 30
CTC AGT GAC AAG GTT CAC AAT GAT GCT CAG AGT TTT GAT TAT GAC CAT 203
Leu Ser Asp Lys Val His Asn Asp Ala Gln Ser Phe Asp Tyr Asp His
35 40 45
GAT GCC TTC TTG GGT GCT GAA GAA GCA AAG ACC TTT GAT CAG CTG ACA 251
Asp Ala Phe Leu Gly Ala Glu Glu Ala Lys Thr Phe Asp Gln Leu Thr
50 55 60 65
CCA GAA GAG AGC AAG GAA AGG CTT GGA AAG ATT GTA AGT AAA ATA GAT 299
Pro Glu Glu Ser Lys Glu Arg Leu Gly Lys Ile Val Ser Lys Ile Asp

GAG GAA AAA GAA GAA

Gly Asp Lys Asp Gly Rhe Val Thr Val Asp Glu Leu Lys Asp Trp Ile
 85 90 95
 AAA TTT GCA CAA AAG CGC TGG ATT TAC GAG GAT GTA GAG CGA CAG TGG 395
 Lys Phe Ala Gln Lys Arg Trp Ile Tyr Glu Asp Val Glu Arg Gln Trp
 100 105 110
 AAG GGG CAT GAC CTC AAT GAG GAC GGC CTC GTT TCC TGG GAG GAG TAT 443
 Lys Gly His Asp Leu Asn Glu Asp Gly Leu Val Ser Trp Glu Glu Tyr
 115 120 125
 AAA AAT GCC ACC TAC GGC TAC GTT TTA GAT GAT CCA GAT CCT GAT GAT 491
 Lys Asn Ala Thr Tyr Gly Tyr Val Leu Asp Asp Pro Asp Pro Asp Asp
 130 135 140 145
 GGA TTT AAC TAT AAA CAG ATG ATG GTT AGA GAT GAG CGG AGG TTT AAA 539
 Gly Phe Asn Tyr Lys Gln Met Met Val Arg Asp Glu Arg Arg Phe Lys
 150 155 160
 ATG GCA GAC AAG GAT GGA GAC CTC ATT GCC ACC AAG GAG GAG TTC ACA 587
 Met Ala Asp Lys Asp Gly Asp Leu Ile Ala Thr Lys Glu Glu Phe Thr
 165 170 175
 GCT TTC CTG CAC CCT GAG GAG TAT GAC TAC ATG AAA GAT ATA GTA GTA 635
 Ala Phe Leu His Pro Glu Glu Tyr Asp Tyr Met Lys Asp Ile Val Val
 180 185 190
 CAG GAA ACA ATG GAA GAT ATA GAT AAG AAT GCT GAT GGT TTC ATT GAT 683
 Gln Glu Thr Met Glu Asp Ile Asp Lys Asn Ala Asp Gly Phe Ile Asp
 195 200 205
 CTA GAA GAG TAT ATT GGT GAC ATG TAC AGC CAT GAT GGG AAT ACT GAT 731
 Leu Glu Glu Tyr Ile Gly Asp Met Tyr Ser His Asp Gly Asn Thr Asp
 210 215 220 225
 GAG CCA GAA TGG GTA AAG ACA GAG CGA GAG CAG TTT GTT GAG TTT CGG 779
 Gly Phe Val Glu Phe Arg

GAT AAG AAC CGT GAT GGG AAG ATG GAC AAG GAA GAG ACC AAA GAC TGG 827
 Asp Lys Asn Arg Asp Gly Lys Met Asp Lys Glu Glu Thr Lys Asp Trp
 245 250 255
 ATC CTT CCC TCA GAC TAT GAT CAT GCA GAG GCA GAA GCC AGG CAC CTG 875
 Ile Leu Pro Ser Asp Tyr Asp His Ala Glu Ala Glu Ala Arg His Leu
 260 265 270
 GTC TAT GAA TCA GAC CAA AAC AAG GAT GGC AAG CTT ACC AAG GAG GAG 923
 Val Tyr Glu Ser Asp Gln Asn Lys Asp Gly Lys Leu Thr Lys Glu Glu
 275 280 285
 ATC GTT GAC AAG TAT GAC TTA TTT GTT GGC AGC CAG GCC ACA GAT TTT 971
 Ile Val Asp Lys Tyr Asp Leu Phe Val Gly Ser Gln Ala Thr Asp Phe
 290 295 300 305
 GGG GAG GCC TTA GTA CGG CAT GAT GAG TTC TGAGCTACGG AGGAACCCT 1020
 Gly Glu Ala Leu Val Arg His Asp Glu Phe
 310 315
 CATTTCCTCA AAAGTAATTT ATTTTACAG CTCTGGITT CACATGAAAT TGTTTGGCT 1080
 ACTGAGACTG TTAACATAAA CTTTTAAGA CATGAAAAGG CGTAATGAAA ACCATCCCGT 1140
 CCCCATTCCT CCTCCTCTCT GAGGGACTGG AGGGAAGCCG TGCTTCTGAG GAACAACTCT 1200
 AATTAGTACA CTGTGTTTG TAGATTTACA CTTTGTATTA TGTATTAACA TGGCGTGT 1260
 ATTTTGTAT TTTTCTCTGG TTGGGAGTAT GATATGAAGG ATCAAGATCC TCAACTCACA 1320
 CATGTAGACA AACATTAGCT CTTTACTCTT TCTCAACCCC TTTTATGATT TTAATAATTC 1380
 TCACTTAACT AATTTTGTA GCCTGAGATC AATAAGAAAT GTTCAGGAGA GAGGAAAGAA 1440
 AAAAAATATA TGCTCCACAA TTTATATTTA GAGAGAGAAC ACTTAGTCTT GCCTGTCAAA 1500
 AAGTCCAACA TTTCATAGGT AGTAGGGGCC ACATATTACA TTCAGTTGCT ATAGGTCCAG 1560
 CAACTGAACC TGCCATTACC TGGGCAAGGA AAGATCCCTT TGCTCTAGGA AAGCTTGGCC 1620
 CAAATTGATT TTCTTCTTTT TCCCCCTGTA GGAAGTACTG TTGGCTAATT TTGTCAGCA 1680
 CAGCTGTGGT GGGAAGAGTT AGGGCCAGTG TCTTGAAAAT CAATCAAGTA GTGAATGTGA 1740
 1800

TGCCTTTTGA AATCACTGTA AATGCCCCCA TCCGGTTCCT CTTCTTCCCA GGTGTGCCAA	1920
GGAATTAATC TTGGTTTCAC TACAATTAAA ATTCACTCCT TTCCAATCAT GTCATTGAAA	1980
GTGCCTTTAA CGAAAGAAAT GGTCACTGAA TGGGAATTCT CTTAAGAAAC CCTGAGATTA	2040
AAAAAAGACT ATTTGGATAA CTTATAGGAA AGCCTAGAAC CTCCCAGTAG AGTGGGGATT	2100
TTTTTCTTCT TCCCTTTCTC TTTTGGACAA TAGTTAAATT AGCAGTATTA GTTATGAGTT	2160
TGGTTGCAGT GTTCTTATCT TGTGGGCTGA TTTCCAAAAA CCACATGCTG CTGAATTTAC	2220
CAGGGATCCT CATACTCAC AATGCAAACC ACTTACTACC AGGCCTTTTT CTGTGTCCAC	2280
TGGAGAGCTT GAGCTCACAC TCAAAGATCA GAGGACCTAC AGAGAGGGCT CTTTGGTTTG	2340
AGGACCATGG CTTACCTTTC CTGCCTTTGA CCCATCACAC CCCATTTCTT CCTCTTTCCC	2400
TCTCCCCGCT GCCAAAAAAA AAAAAAAAAG GAAACGTTTA TCATGAATCA ACAGGGTTTC	2460
AGTCCTTATC AAAGAGAGAT GTGGAAAGAG CTAAAGAAAC CACCCTTTGT TCCCAACTCC	2520
ACTTTACCCA TATTTTATGC AACACAAACA CTGTCTTTT GGGTCCCTTT CTTACAGATG	2580
GACCTCTTGA GAAGAATTAT CGTATTCCAC GTTTTTAGCC CTCAGGTTAC CAAGATAAAT	2640
ATATGTATAT ATAACCTTTA TTATTGCTAT ATCTTTGTGG ATAATACATT CAGGTGGTGC	2700
TGGGTGATTT ATTATAATCT GAACCTAGGT ATATCCTTTG GTCTTCCACA GTCATGTTGA	2760
GGTGGGCTCC CTGGTATGGT AAAAAGCCAG GTATAATGTA ACTTCACCCC AGCCTTTGTA	2820
CTAAGCTCTT GATAGTGGAT ATACTCTTTT AAGTTTAGCC CCAATATAGG GTAATGAAAA	2880
TTTCCTGCCC TCTGGGTTCC CCATTTTAC TATTAAGAAG ACCAGTGATA ATTTAATAAT	2940
GCCACCAACT CTGGCTTAGT TAAGTGAGAG TGTGAACTGT GTGGCAAGAG AGCCTCACAC	3000
CTCACTAGGT GCAGAGAGCC CAGGCCTTAT GTTAAATCA TGCACCTGAA AAGCAAACCT	3060
TAATCTGCAA AGACAGCAGC AAGCATTATA CGGTCATCTT GAATGATCCC TTTGAAATTT	3120
TTTTTTTGTT TGTTTGTTTA AATCAAGCCT GAGGCTGGTG AACAGTAGCT ACACACCCAT	3180
ATTGTGTGTT CTGTGAATGC TAGCTTTCTT GAATTTGGAT ATTGGTTATT TTTTATAGAG	3240
TGTAAACCAA GTTTTATATT CTGCAATGCG AACAGGTACC TATCTGTTTC TAAATAAAAC	3300
TGTTTACATT C	3311

Sequence No.: 21

Strandedness: Double

Topology: Linear

Sequence kind: cDNA to mRNA

Original source:

Organism species: *Homo sapiens*

Cell kind: Stomach cancer

Clone name: HP00876

Sequence characteristics:

Code representing characteristics: CDS

Existence site: 147.. 623

Characterization method: E

Sequence description

ACTGGAGACA CTGAAGAAGG CAGGGGCCCT TAGAGTCTTG GTTGCCAAAC AGATTTGCAG	60
ATCAAGGAGA ACCCAGGAGT TTCAAAGAAG CGCTAGTAAG GTCTCTGAGA TCCTTGCACT	120
AGCTACATCC TCAGGGTAGG AGGAAG ATG GCT TCC AGA AGC ATG CGG CTG CTC	173
Met Ala Ser Arg Ser Met Arg Leu Leu	
1 5	
CTA TTG CTG AGC TGC CTG GCC AAA ACA GGA GTC CTG GGT GAT ATC ATC	221
Leu Leu Leu Ser Cys Leu Ala Lys Thr Gly Val Leu Gly Asp Ile Ile	
10 15 20 25	
ATG AGA CCC AGC TGT GCT CCT GGA TGG TTT TAC CAC AAG TCC AAT TGC	269
Met Arg Pro Ser Cys Ala Pro Gly Trp Phe Tyr His Lys Ser Asn Cys	
30 35 40	
TAT GGT TAC TTC AGG AAG CTG AGG AAC TGG TCT GAT GCC GAG CTC GAG	317
Tyr Gly Tyr Phe Arg Lys Leu Arg Asn Trp Ser Asp Ala Glu Leu Glu	
45 50 55	
TGT CAG TCT TAC GGA AAC GGA GCC CAC CTG GCA TCT ATC CTG AGT TTA	365

AAG GAA GCC AGC ACC ATA GCA GAG TAC ATA AGT GGC TAT CAG AGA AGC 413
 Lys Glu Ala Ser Thr Ile Ala Glu Tyr Ile Ser Gly Tyr Gln Arg Ser
 75 80 85
 CAG CCG ATA TGG ATT GGC CTG CAC GAC CCA CAG AAG AGG CAG CAG TGG 461
 Gln Pro Ile Trp Ile Gly Leu His Asp Pro Gln Lys Arg Gln Gln Trp
 90 95 100 105
 CAG TGG ATT GAT GGG GCC ATG TAT CTG TAC AGA TCC TGG TCT GGC AAG 509
 Gln Trp Ile Asp Gly Ala Met Tyr Leu Tyr Arg Ser Trp Ser Gly Lys
 110 115 120
 TCC ATG GGT GGG AAC AAG CAC TGT GCT GAG ATG AGC TCC AAT AAC AAC 557
 Ser Met Gly Gly Asn Lys His Cys Ala Glu Met Ser Ser Asn Asn Asn
 125 130 135
 TTT TTA ACT TGG AGC AGC AAC GAA TGC AAC AAG CGC CAA CAC TTC CTG 605
 Phe Leu Thr Trp Ser Ser Asn Glu Cys Asn Lys Arg Gln His Phe Leu
 140 145 150
 TGC AAG TAC CGA CCA TAGAGCAAGA ATCAAGATTC TGCTAACTCC 650
 Cys Lys Tyr Arg Pro
 155
 TGCACAGCCC CGTCCTCTTC CTTTCTGCTA GCCTGGCTAA ATCTGCTCAT TATTCAGAG 710
 GGGAAACCTA GCAAACCTAAG AGTGATAAGG GCCCTACTAC ACTGGCTTTT TTAGGCTTAG 770
 AGACAGAAAC TTTAGCATTG GCCCAGTAGT GGCTTCTAGC TCTAAATGTT TGCCCCGCCA 830
 TCCCTTTCCA CAGTATCCTT CTTCCCTCCT CCCCTGTCTC TGGCTGTCTC GAGCAGTCTA 890
 GAAGAGTGCA TCTCCAGCCT ATGAAACAGC TGGGTCTTTG GCCATAAGAA GTAAAGATTT 950
 GAAGACAGAA GGAAGAAACT CAGGAGTAAG CTCTAGCCC CCTTCAGCTT CTACACCCTT 1010
 CTGCCCTCTC TCCATTGCCT GCACCCACCC CCAGCCACTC AACTCCTGCT TGTTTTTCCT 1070
 TTGGCCATGG GAAGGTTTAC CAGTAGAATC CTTGCTAGGT TGATGTGGGC CATACATTCC 1130
 TTTAATAAAC CATTGTGTAC AT 1152

Sequence length: 1749

Sequence type: Nucleic acid

Strandedness: Double

Topology: Linear

Sequence kind: cDNA to mRNA

Original source:

Organism species: *Homo sapiens*

Cell kind: Liver

Clone name: HP01134

Sequence characteristics:

Code representing characteristics: CDS

Existence site: 117.. 1247

Characterization method: E

Sequence description

AATCACAGCA GTNCCGACGT CGTGGGTGTT TGGTGTGAGG CTGCGAGCCG CCGCCGCCAC 60
CACTGCCACC ACGGTCGCCT GCCACAGGTG TCTGCAATTG AACTCCAAGG TGCAGA ATG 119

Met

1

GTT TGG AAA GTA GCT GTA TTC CTC AGT GTG GCC CTG GGC ATT GGT GCC 167
Val Trp Lys Val Ala Val Phe Leu Ser Val Ala Leu Gly Ile Gly Ala

5

10

15

GTT CCT ATA GAT GAT CCT GAA GAT GGA GGC AAG CAC TGG GTG GTG ATC 215
Val Pro Ile Asp Asp Pro Glu Asp Gly Gly Lys His Trp Val Val Ile

20

25

30

GTG GCA GGT TCA AAT GGC TGG TAT AAT TAT AGG CAC CAG GCA GAC GCG 263
Val Ala Gly Ser Asn Gly Trp Tyr Asn Tyr Arg His Gln Ala Asp Ala

35

40

45

98

50	55	60	65	
ATC GTT GTG ATG ATG TAC GAT GAC ATT GCT TAC TCT GAA GAC AAT CCC				359
Ile Val Val Met Met Tyr Asp Asp Ile Ala Tyr Ser Glu Asp Asn Pro				
70	75	80		
ACT CCA GGA ATT GTG ATC AAC AGG CCC AAT GGC ACA GAT GTC TAT CAG				407
Thr Pro Gly Ile Val Ile Asn Arg Pro Asn Gly Thr Asp Val Tyr Gln				
85	90	95		
GGA GTC CCG AAG GAC TAC ACT GGA GAG GAT GTT ACC CCA CAA AAT TTC				455
Gly Val Pro Lys Asp Tyr Thr Gly Glu Asp Val Thr Pro Gln Asn Phe				
100	105	110		
CTT GCT GTG TTG AGA GGC GAT GCA GAA GCA GTG AAG GGC ATA GGA TCC				503
Leu Ala Val Leu Arg Gly Asp Ala Glu Ala Val Lys Gly Ile Gly Ser				
115	120	125		
GGC AAA GTC CTG AAG AGT GGC CCC CAG GAT CAC GTG TTC ATT TAC TTC				551
Gly Lys Val Leu Lys Ser Gly Pro Gln Asp His Val Phe Ile Tyr Phe				
130	135	140	145	
ACT GAC CAT GGA TCT ACT GGA ATA CTG GTT TTT CCC AAT GAA GAT CTT				599
Thr Asp His Gly Ser Thr Gly Ile Leu Val Phe Pro Asn Glu Asp Leu				
150	155	160		
CAT GTA AAG GAC CTG AAT GAG ACC ATC CAT TAC ATG TAC AAA CAC AAA				647
His Val Lys Asp Leu Asn Glu Thr Ile His Tyr Met Tyr Lys His Lys				
165	170	175		
ATG TAC CGA AAG ATG GTG TTC TAC ATT GAA GCC TGT GAG TCT GGG TCC				695
Met Tyr Arg Lys Met Val Phe Tyr Ile Glu Ala Cys Glu Ser Gly Ser				
180	185	190		
ATG ATG AAC CAC CTG CCG GAT AAC ATC AAT GTT TAT GCA ACT ACT GCT				743
Met Met Asn His Leu Pro Asp Asn Ile Asn Val Tyr Ala Thr Thr Ala				

... AAC CCG AGA GAG ...

Ala Asn Pro Arg Glu Ser Ser Tyr Ala Cys Tyr Tyr Asp Glu Lys Arg
 210 215 220 225
 TCC ACG TAC CTG GGG GAC TGG TAC AGC GTC AAC TGG ATG GAA GAC TCG 839
 Ser Thr Tyr Leu Gly Asp Trp Tyr Ser Val Asn Trp Met Glu Asp Ser
 230 235 240
 GAC GTG GAA GAT CTG ACT AAA GAG ACC CTG CAC AAG CAG TAC CAC CTG 887
 Asn Val Glu Asp Leu Thr Lys Glu Thr Leu His Lys Gln Tyr His Leu
 245 250 255
 GTA AAA TCG CAC ACC AAC ACC AGC CAC GTC ATG CAG TAT GGA AAC AAA 935
 Val Lys Ser His Thr Asn Thr Ser His Val Met Gln Tyr Gly Asn Lys
 260 265 270
 ACA ATC TCC ACC ATG AAA GTG ATG CAG TTT CAG GGT ATG AAA CGC AAA 983
 Thr Ile Ser Thr Met Lys Val Met Gln Phe Gln Gly Met Lys Arg Lys
 275 280 285
 GCC AGT TCT CCC GTC CCC CTA CCT CCA GTC ACA CAC CTT GAC CTC ACC 1031
 Ala Ser Ser Pro Val Pro Leu Pro Pro Val Thr His Leu Asp Leu Thr
 290 295 300 305
 CCC AGC CCT GAT GTG CCT CTC ACC ATC ATG AAA AGG AAA CTG ATG AAC 1079
 Pro Ser Pro Asp Val Pro Leu Thr Ile Met Lys Arg Lys Leu Met Asn
 310 315 320
 ACC AAT GAT CTG GAG GAG TCC AGG CAG CTC ACG GAG GAG ATC CAG CGG 1127
 Thr Asn Asp Leu Glu Glu Ser Arg Gln Leu Thr Glu Glu Ile Gln Arg
 325 330 335
 CAT CTG GAT TAC GAG TAT GCG TTG AGA CAT TTG TAC GTG CTG GTC AAC 1175
 His Leu Asp Tyr Glu Tyr Ala Leu Arg His Leu Tyr Val Leu Val Asn
 340 345 350
 CTT TGT GAG AAG CCG TAT CCG CTT CAC AGG ATA AAA TTG TCC ATG GAC 1223

100

CAC GTG TGC CTT GGT CAC TAC TGAAGAGCTG CCTCCTGGAA GCTTTT 1270
His Val Cys Leu Gly His Tyr
370 375
CCAAGTGTGA GCGCCCCACC GACTGTGTGC TGATCAGAGA CTGGAGAGGT GGAGTGAGAA 1330
GTCTCCGCTG CTCGGGGCCCT CCTGGGGAGC CCCCCTCCA GGGCTCGCTC CAGGACCTTC 1390
TTCACAAGAT GACTTGCTCG CTGTTACCTG CTTCCCCAGT CTTTCTGAA AACTACAAA 1450
TTAGGGTGGG AAAAGCTCTG TATTGAGAAG GGTCAATTTT GCTTCTAGG AGGTTTGTG 1510
TTTTCCTGT TAGTTTGTAG GAGCAGGAAG CTCATGGGGG CTTCTGTAGC CTTCTGAAA 1570
AGGAGTCTTT ATTCTGAGAA TTTGAAGCTG AAACCTCTTT AAATCTTCAG AATGATTTTA 1630
TTGAAGAGGG CCGCAAGCCC CAAATGAAA ACTGTTTTTA GAAAATATGA TGATTTTGA 1690
TTGCTTTTGT ATTTAATTCT GCAGGTGTTT AACTCTTAAA AAATAAAGAT TTATAACAG 1749

Sequence No.: 23

Sequence length: 988

Sequence type: Nucleic acid

Strandedness: Double

Topology: Linear

Sequence kind: cDNA to mRNA

Original source:

Organism species: *Homo sapiens*

Cell kind: Epidermoid carcinoma

Cell line: KB

Clone name: HP10029

Sequence characteristics:

Code representing characteristics: CDS

Existence site: 9.. 530

Characterization method: E

101

Met Ala Ala Pro Ser Gly Gly Trp Asn Gly Val Arg Ala Ser

1	5	10	
TTG TGG GCC GCG CTG CTC CTA GGG GCC GTG GCG CTG AGG CCG GCG GAG			98
Leu Trp Ala Ala Leu Leu Leu Gly Ala Val Ala Leu Arg Pro Ala Glu			
15	20	25	30
CGG GTG TCC GAG CCC ACC ACC GTG GCG TTT GAC GTG CCG CCC GGC GGC			146
Ala Val Ser Glu Pro Thr Thr Val Ala Phe Asp Val Arg Pro Gly Gly			
35	40	45	
GTC GTG CAT TCC TTC TCC CAT AAC GTG GGC CCG GGG GAC AAA TAT ACG			194
Val Val His Ser Phe Ser His Asn Val Gly Pro Gly Asp Lys Tyr Thr			
50	55	60	
TGT ATG TTC ACT TAC GCC TCT CAA GGA GGG ACC AAT GAG CAA TGG CAG			242
Cys Met Phe Thr Tyr Ala Ser Gln Gly Gly Thr Asn Glu Gln Trp Gln			
65	70	75	
ATG AGT CTG GGG ACC AGC GAA GAC CAC CAG CAC TTC ACC TGC ACC ATC			290
Met Ser Leu Gly Thr Ser Glu Asp His Gln His Phe Thr Cys Thr Ile			
80	85	90	
TGG AGG CCC CAG GGG AAG TCC TAT CTG TAC TTC ACA CAG TTC AAG GCA			338
Trp Arg Pro Gln Gly Lys Ser Tyr Leu Tyr Phe Thr Gln Phe Lys Ala			
95	100	105	110
GAG GTG CCG GGC GCT GAG ATT GAG TAC GCC ATG GCC TAC TCT AAA GCC			386
Glu Val Arg Gly Ala Glu Ile Glu Tyr Ala Met Ala Tyr Ser Lys Ala			
115	120	125	
GCA TTT GAA AGG GAA AGT GAT GTC CCT CTG AAA ACT GAG GAA TTT GAA			434
Ala Phe Glu Arg Glu Ser Asp Val Pro Leu Lys Thr Glu Glu Phe Glu			
130	135	140	
GTG ACC AAA ACA GCA GTG GCT CAC AGG CCC GGG GCA TTC AAA GCT GAG			482

102

CTG TCC AAG CTG GTG ATT GTG GCC AAG GCA TCG CGC ACT GAG CTG 527
Leu Ser Lys Leu Val Ile Val Ala Lys Ala Ser Arg Thr Glu Leu
160 165 170
TGA CCAGCAGCCC TGTTGCGGGT GGCACCTTCT CATCTCCGGT GAAGCTGAAG 580
GGGCCTGTGG CCCTGAAAGG GCCAGCACAT CACTGGTTTT CTAGGAGGGA CTCTTAAGTT 640
TTCTACCTGG GCTGACGTTG CCTTGTCGG AGGGGCTTGC AGGGTGGCTG AAGCCCTGGG 700
GCAGAGAACA GAGGGTCCAG GGCCCTCCTG GCTCCCAACA GCTTCTCAGT TCCCACTTCC 760
TGCTGAGCTC TTCTGGACTC AGGATCGCAG ATCCGGGGCA CAAAGAGGGT GGGGAACATG 820
GGGGCTATGC TGGGGAAAGC AGCCATGCTC CCCCCGACCT CCAGCCGAGC ATCCTTCATG 880
AGCCTGCAGA ACTGCTTTCC TATGTTTACC CAGGGGACCT CCTTTCAGAT GAACTGGGAA 940
GAGATGAAAT GTTTTTTCAT ATTTAAATAA ATAAGAACAT TAAAAAGC 988

Sequence No.: 24

Sequence length: 390

Sequence type: Nucleic acid

Strandedness: Double

Topology: Linear

Sequence kind: cDNA to mRNA

Original source:

Organism species: *Homo sapiens*

Cell kind: Epidermoid carcinoma

Cell line: KB

Clone name: HP10189

Sequence characteristics:

Code representing characteristics: CDS

Existence site: 102.. 323

Characterization method: E

103

GGCTTTGGGC GGAAC TGGCT TTGTTGACCG GGAGAAACGA G ATG GGG GTG AAG CTG 116
 Met Gly Val Lys Leu
 1 5
 GAG ATA TTT CGG ATG ATA ATC TAC CTC ACT TTC CCT GTG GCT ATG TTC 164
 Glu Ile Phe Arg Met Ile Ile Tyr Leu Thr Phe Pro Val Ala Met Phe
 10 15 20
 TGG GAT TCG AAT GAT CCG CAG TGG TTT GAG GAC GAT GTC ATA CAG CGC 212
 Trp Val Ser Asn Gln Ala Glu Trp Phe Glu Asp Asp Val Ile Gln Arg
 25 30 35
 AAG AGG GAG CTG TGG CCA CCT GAG AAG CTT CAA GAG ATA GAG GAA TTC 260
 Lys Arg Glu Leu Trp Pro Pro Glu Lys Leu Gln Glu Ile Glu Glu Phe
 40 45 50
 AAA GAG AGG TTA CGG AAG CGG CGG GAG GAG AAG CTC CTT CGC GAC GCC 308
 Lys Glu Arg Leu Arg Lys Arg Arg Glu Glu Lys Leu Leu Arg Asp Ala
 55 60 65
 CAG CAG AAC TCC TGAGGCCTCC AAGTGGGAGT CCTAGCCCCT 350
 Gln Gln Asn Ser
 70
 CCCCTGATGA AATATACATA TACTCAGTTC CTTGTTATTC 390

Sequence No.: 25

Sequence length: 4667

Sequence type: Nucleic acid

Strandedness: Double

Topology: Linear

Sequence kind: cDNA to mRNA

Original source:

Accession

Accession Kind: Lymphoma

Cell line: U937

Clone name: HP10269

Sequence characteristics:

Code representing characteristics: CDS

Existence site: 754.. 4272

Characterization method: E

Sequence description

CATTAGTTA CTCTGCTCAT TTCTCTTAAG CTTTCCTTGG ATGAGTTGAG CTTTGAATCC 60
 TTCCTGATGA ACCTTGCCTT TTAAGGATCC TCCAAATGCC CCAAGAAGCT GGGATTTTTC 120
 ATTTTTTTTT TCACTGGGGA GGGGAATGGT GCTTTCCAGG GTCCTGGATG TTTGAGTCTT 180
 CTCACCTTCC AGCCCGGTGA TATGTCTGGA GCTTTAACTC TCTATATAAG CCCTAATCTT 240
 TGTGTTCTCT GCCTGATCTT CTGTCTGGGG TGGTCCAGGT CACAAGAAGA AGCTGACCCC 300
 TGCTGGCTTT GGGAAAATGC TGAGTTCATT GCCTGGCACA AATGCAAGGG CCCTTCCCCA 360
 CCCTGTGAAT TCTGGTCTCT GATGATCACT TACATGTGCC TTGTGCTTTC TGTTTGAGGG 420
 GCCCCTTGCA GCCCCCACAG GCAGGTGGGC ATTGTGGAGC TCACTACAAG AACTCTGGGA 480
 CCGACCGACC AACCCACTTG CCCAGTCCCG TCCTGGGAGG TGGGGGTGCA GTGACGACAG 540
 ATGGGTGTGA CGGCTGCCAG ATTCCTGAGA CCCGCCCTGC GGTGGGGCTA CACCCAGCCA 600
 GGGAGTCTCC AGAGGTGAGG CTGTTGTTTA AAAACCTGGA GCCGGGAGGG GAGACCCCCA 660
 CATTCAAGAG GAGCTTTCAG GCGATCTGGA GAAAGAACGG CAGAACACAC AGCAAGGAAA 720
 GGTCTTTTCT GGGGATCACC CCATTGGCTG AAG ATG AGA CCA TTC TTC CTC TTG 774

Met Arg Pro Phe Phe Leu Leu

1

5

TGT TTT GCC CTG CCT GGC CTC CTG CAT GCC CAA CAA GCC TGC TCC CGT 822
 Cys Phe Ala Leu Pro Gly Leu Leu His Ala Gln Gln Ala Cys Ser Arg

10

15

20

GGG GCC TGC TAT CCA CCT GTT GGG GAC CTG CTT GTT GGG AGG ACC CGG 870
 Gly Ala Cys Tyr Pro Pro Val Gly Asp Leu Leu Val Gly Arg Thr Arg

105

Phe	Leu	Arg	Ala	Ser	Ser	Thr	Cys	Gly	Leu	Thr	Lys	Pro	Glu	Thr	Tyr		
40					45					50					55		
TGC	ACC	CAG	TAT	GGC	GAG	TGG	CAG	ATG	AAA	TGC	TGC	AAG	TGT	GAC	TCC	966	
Cys	Thr	Gln	Tyr	Gly	Glu	Trp	Gln	Met	Lys	Cys	Cys	Lys	Cys	Asp	Ser		
				60					65						70		
AGG	CAG	CCT	CAC	AAC	TAC	TAC	AGT	CAC	CGA	GTA	GAG	AAT	GTG	GCT	TCA	1014	
Arg	Gln	Phe	His	Asp	Tyr	Tyr	Ser	His	Arg	Val	Glu	Asn	Val	Ala	Ser		
				75					80						85		
TCC	TCC	GGC	CCC	ATG	CGC	TGG	TGG	CAG	TCC	CAG	AAT	GAT	GTG	AAC	CCT	1062	
Ser	Ser	Gly	Pro	Met	Arg	Trp	Trp	Gln	Ser	Gln	Asn	Asp	Val	Asn	Pro		
				90					95						100		
GTC	TCT	CTG	CAG	CTG	GAC	CTG	GAC	AGG	AGA	TTC	CAG	CTT	CAA	GAA	GTC	1110	
Val	Ser	Leu	Gln	Leu	Asp	Leu	Asp	Arg	Arg	Phe	Gln	Leu	Gln	Glu	Val		
				105					110						115		
ATG	ATG	GAG	TTC	CAG	GGG	CCC	ATG	CCT	GCC	GGC	ATG	CTG	ATT	GAG	CGC	1158	
Met	Met	Glu	Phe	Gln	Gly	Pro	Met	Pro	Ala	Gly	Met	Leu	Ile	Glu	Arg		
120					125					130					135		
TCC	TCA	GAC	TTC	GGT	AAG	ACC	TGG	CGA	GTG	TAC	CAG	TAC	CTG	GCT	GCC	1206	
Ser	Ser	Asp	Phe	Gly	Lys	Thr	Trp	Arg	Val	Tyr	Gln	Tyr	Leu	Ala	Ala		
				140					145						150		
GAC	TGC	ACC	TCC	ACC	TTC	CCT	CGG	GTG	CGC	CAG	GGT	CGG	CCT	CAG	AGC	1254	
Asp	Cys	Thr	Ser	Thr	Phe	Pro	Arg	Val	Arg	Gln	Gly	Arg	Pro	Gln	Ser		
				155					160						165		
TGG	CAG	GAT	GTT	CGG	TGC	CAG	TCC	CTG	CCT	CAG	AGG	CCT	AAT	GCA	CGC	1302	
Trp	Gln	Asp	Val	Arg	Cys	Gln	Ser	Leu	Pro	Gln	Arg	Pro	Asn	Ala	Arg		
				170					175						180		
CTA	AAT	GGG	GGG	AAG	GTC	CAA	CTT	AAC	CTT	ATG	GAT	TTA	GTG	TCT	GGG	1350	

ATT CCA GCA ACT CAA AGT CAA AAA ATT CAA GAG GTG GGG GAG ATC ACA	1398
Ile Pro Ala Thr Gln Ser Gln Lys Ile Gln Glu Val Gly Glu Ile Thr	
200 205 210 215	
AAC TTG AGA GTC AAT TTC ACC AGG CTG GCC CCT GTG CCC CAA AGG GGC	1446
Asn Leu Arg Val Asn Phe Thr Arg Leu Ala Pro Val Pro Gln Arg Gly	
220 225 230	
TAC CAC CCT CCC AGC GCC TAC TAT GCT GTG TCC CAG CTC CGT CTG CAG	1494
Tyr His Pro Pro Ser Ala Tyr Tyr Ala Val Ser Gln Leu Arg Leu Gln	
235 240 245	
GGG AGC TGC TTC TGT CAC GGC CAT GCT GAT CGC TGC GCA CCC AAG CCT	1542
Gly Ser Cys Phe Cys His Gly His Ala Asp Arg Cys Ala Pro Lys Pro	
250 255 260	
GGG GCC TCT GCA GGC CCC TCC ACC GCT GTG CAG GTC CAC GAT GTC TGT	1590
Gly Ala Ser Ala Gly Pro Ser Thr Ala Val Gln Val His Asp Val Cys	
265 270 275	
GTC TGC CAG CAC AAC ACT GCC GGC CCA AAT TGT GAG CGC TGT GCA CCC	1638
Val Cys Gln His Asn Thr Ala Gly Pro Asn Cys Glu Arg Cys Ala Pro	
280 285 290 295	
TTC TAC AAC AAC CGG CCC TGG AGA CCG GCG GAG GGC CAG GAC GCC CAT	1686
Phe Tyr Asn Asn Arg Pro Trp Arg Pro Ala Glu Gly Gln Asp Ala His	
300 305 310	
GAA TGC CAA AGG TGC GAC TGC AAT GGG CAC TCA GAG ACA TGT CAC TTT	1734
Glu Cys Gln Arg Cys Asp Cys Asn Gly His Ser Glu Thr Cys His Phe	
315 320 325	
GAC CCC GCT GTG TTT GCC GCC AGC CAG GGG GCA TAT GGA GGT GTG TGT	1782
Asp Pro Ala Val Phe Ala Ala Ser Gln Gly Ala Tyr Gly Gly Val Cys	
330 335 340	

$$As_2O_3 + AsH_3 \rightleftharpoons As_2O_5 + H_2O$$

107

345	350	355	
CTG CAC TAT TTC CGG AAC CGG CGC CCG GGA GCT TCC ATT CAG GAG ACC			1878
Leu His Tyr Phe Arg Asn Arg Arg Pro Gly Ala Ser Ile Gln Glu Thr			
360	365	370	375
TGC ATC TCC TGC GAG TGT GAT CCG GAT GGG GCA GTG CCA GGG GCT CCC			1926
Cys Ile Ser Cys Glu Cys Asp Pro Asp Gly Ala Val Pro Gly Ala Pro			
380	385	390	
TGT GAC CCA GTG ACC GGG CAG TGT GTG TGC AAG GAG CAT GTG CAG GGA			1974
Cys Asp Pro Val Thr Gly Gln Cys Val Cys Lys Glu His Val Gln Gly			
395	400	405	
GAG CGC TGT GAC CTA TGC AAG CCG GGC TTC ACT GGA CTC ACC TAC GCC			2022
Glu Arg Cys Asp Leu Cys Lys Pro Gly Phe Thr Gly Leu Thr Tyr Ala			
410	415	420	
AAC CCG CAG GGC TGC CAC CGC TGT GAC TGC AAC ATC CTG GGG TCC CGG			2070
Asn Pro Gln Gly Cys His Arg Cys Asp Cys Asn Ile Leu Gly Ser Arg			
425	430	435	
AGG GAC ATG CCG TGT GAC GAG GAG AGT GGG CGC TGC CTT TGT CTG CCC			2118
Arg Asp Met Pro Cys Asp Glu Glu Ser Gly Arg Cys Leu Cys Leu Pro			
440	445	450	455
AAC GTG GTG GGT CCC AAA TGT GAC CAG TGT GCT CCC TAC CAC TGG AAG			2166
Asn Val Val Gly Pro Lys Cys Asp Gln Cys Ala Pro Tyr His Trp Lys			
460	465	470	
CTG GCC AGT GGC CAG GGC TGT GAA CCG TGT GCC TGC GAC CCG CAC AAC			2214
Leu Ala Ser Gly Gln Gly Cys Glu Pro Cys Ala Cys Asp Pro His Asn			
475	480	485	
TCC CTC AGC CCA CAG TGC AAC CAG TTC ACA GGG CAG TGC CCC TGT CGG			2262
Ser Leu Ser Pro Gln Cys Asn Gln Phe Thr Gly Gln Cys Pro Cys Arg			

Glu Gly Phe Gly Gly Leu Met Cys Ser Ala Ala Ala Ile Arg Gln Cys
 505 510 515
 CCA GAC CGG ACC TAT GGA GAC GTG GCC ACA GGA TGC CGA GCC TGT GAC 2358
 Pro Asp Arg Thr Tyr Gly Asp Val Ala Thr Gly Cys Arg Ala Cys Asp
 520 525 530 535
 TGT GAT TTC CGG GGA ACA GAG GGC CCG GGC TGC GAC AAG GCA TCA GGC 2406
 Cys Asp Phe Arg Gly Thr Glu Gly Pro Gly Cys Asp Lys Ala Ser Gly
 540 545 550
 CGC TGC CTC TGC CGC CCT GGC TTG ACC GGG CCC CGC TGT GAC CAG TGC 2454
 Arg Cys Leu Cys Arg Pro Gly Leu Thr Gly Pro Arg Cys Asp Gln Cys
 555 560 565
 CAG CGA GGC TAC TGC AAT CGC TAC CCG GTG TGC GTG GCC TGC CAC CCT 2502
 Gln Arg Gly Tyr Cys Asn Arg Tyr Pro Val Cys Val Ala Cys His Pro
 570 575 580
 TGC TTC CAG ACC TAT GAT GCG GAC CTC CGG GAG CAG GCC CTG CGC TTT 2550
 Cys Phe Gln Thr Tyr Asp Ala Asp Leu Arg Glu Gln Ala Leu Arg Phe
 585 590 595
 GGT AGA CTC CGC AAT GCC ACC GCC AGC CTG TGG TCA GGG CCT GGG CTG 2598
 Gly Arg Leu Arg Asn Ala Thr Ala Ser Leu Trp Ser Gly Pro Gly Leu
 600 605 610 615
 GAG GAC CGT GGC CTG GCC TCC CGG ATC CTA GAT GCA AAG AGT AAG ATT 2646
 Glu Asp Arg Gly Leu Ala Ser Arg Ile Leu Asp Ala Lys Ser Lys Ile
 620 625 630
 GAG CAG ATC CGA GCA GTT CTC AGC AGC CCC GCA GTC ACA GAG CAG GAG 2694
 Glu Gln Ile Arg Ala Val Leu Ser Ser Pro Ala Val Thr Glu Gln Glu
 635 640 645
 GTG GCT CAG GTG GCC AGT GCC ATC CTC TCC CTC AGG CGA ACT CTC CAG 2742

109

GGC CTG CAG CTG GAT CTG CCC CTG GAG GAG GAG ACG TTG TCC CTT CCG 2790
 Gly Leu Gln Leu Asp Leu Pro Leu Glu Glu Glu Thr Leu Ser Leu Pro
 665 670 675
 AGA GAC CTG GAG AGT CTT GAC AGA AGC TTC AAT GGT CTC CTT ACT ATG 2838
 Arg Asp Leu Glu Ser Leu Asp Arg Ser Phe Asn Gly Leu Leu Thr Met
 680 685 690 695
 TAT CAG AGG AAC ACC CAG CAG TTT GAA AAA ATA AGC AGT GCT GAT CCT 2886
 Tyr Gln Arg Lys Arg Glu Gln Phe Glu Lys Ile Ser Ser Ala Asp Pro
 700 705 710
 TCA GGA GCC TTC CGG ATG CTG AGC ACA GCC TAC GAG CAG TCA GCC CAG 2934
 Ser Gly Ala Phe Arg Met Leu Ser Thr Ala Tyr Glu Gln Ser Ala Gln
 715 720 725
 GCT GCT CAG CAG GTC TCC GAC AGC TCG CGC CTT TTG GAC CAG CTC AGG 2982
 Ala Ala Gln Gln Val Ser Asp Ser Ser Arg Leu Leu Asp Gln Leu Arg
 730 735 740
 GAC AGC CGG AGA GAG GCA GAG AGG CTG GTG CGG CAG GCG GGA GGA GGA 3030
 Asp Ser Arg Arg Glu Ala Glu Arg Leu Val Arg Gln Ala Gly Gly Gly
 745 750 755
 GGA GGC ACC GGC AGC CCC AAG CTT GTG GCC CTG AGG CTG GAG ATG TCT 3078
 Gly Gly Thr Gly Ser Pro Lys Leu Val Ala Leu Arg Leu Glu Met Ser
 760 765 770 775
 TCG TTG CCT GAC CTG ACA CCC ACC TTC AAC AAG CTC TGT GGC AAC TCC 3126
 Ser Leu Pro Asp Leu Thr Pro Thr Phe Asn Lys Leu Cys Gly Asn Ser
 780 785 790
 AGG CAG ATG GCT TGC ACC CCA ATA TCA TGC CCT GGT GAG CTA TGT CCC 3174
 Arg Gln Met Ala Cys Thr Pro Ile Ser Cys Pro Gly Glu Leu Cys Pro
 795 800 805

... ASP AND GLY THR ALA LYS GLY SER ALA ...

110

810	815	820	
AGG GCC GGT GGG GCC TTC TTG ATG GCG GGG CAG GTG GCT GAG CAG CTG			3270
Arg Ala Gly Gly Ala Phe Leu Met Ala Gly Gln Val Ala Glu Gln Leu			
825	830	835	
CGG GGC TTC AAT GCC CAG CTC CAG CGG ACC AGG CAG ATG ATT AGG GCA			3318
Arg Gly Phe Asn Ala Gln Leu Gln Arg Thr Arg Gln Met Ile Arg Ala			
840	845	850	855
GCC GAG GAA TCT GCC TCA CAG ATT CAA TCC AGT GCC CAG CGC TTG GAG			3366
Ala Glu Glu Ser Ala Ser Gln Ile Gln Ser Ser Ala Gln Arg Leu Glu			
860	865	870	
ACC CAG GTG AGC GCC AGC CGC TCC CAG ATG GAG GAA GAT GTC AGA CGC			3414
Thr Gln Val Ser Ala Ser Arg Ser Gln Met Glu Glu Asp Val Arg Arg			
875	880	885	
ACA CGG CTC CTA ATC CAG CAG GTC CGG GAC TTC CTA ACA GAC CCC GAC			3462
Thr Arg Leu Leu Ile Gln Gln Val Arg Asp Phe Leu Thr Asp Pro Asp			
890	895	900	
ACT GAT GCA GCC ACT ATC CAG GAG GTC AGC GAG GCC GTG CTG GCC CTG			3510
Thr Asp Ala Ala Thr Ile Gln Glu Val Ser Glu Ala Val Leu Ala Leu			
905	910	915	
TGG CTG CCC ACA GAC TCA GCT ACT GTT CTG CAG AAG ATG AAT GAG ATC			3558
Trp Leu Pro Thr Asp Ser Ala Thr Val Leu Gln Lys Met Asn Glu Ile			
920	925	930	935
CAG GCC ATT GCA GCC AGG CTC CCC AAC GTG GAC TTG GTG CTG TCC CAG			3606
Gln Ala Ile Ala Ala Arg Leu Pro Asn Val Asp Leu Val Leu Ser Gln			
940	945	950	
ACC AAG CAG GAC ATT GCG CGT GCC CGC CGG TTG CAG GCT GAG GCT GAG			3654
Thr Lys Gln Asp Ile Ala Arg Ala Arg Leu Gln Ala Glu Ala Glu			

111

Glu Ala Arg Ser Arg Ala His Ala Val Glu Gly Gln Val Glu Asp Val
 970 975 980
 GTT GGG AAC CTG CGG CAG GGG ACA GTG GCA CTG CAG GAA GCT CAG GAC 3750
 Val Gly Asn Leu Arg Gln Gly Thr Val Ala Leu Gln Glu Ala Gln Asp
 985 990 995
 ACC ATG CAA GGC ACC AGC CGC TCC CTT CGG CTT ATC CAG GAC AGG GTT 3798
 Thr Met Gln Gly Thr Ser Arg Ser Leu Arg Leu Ile Gln Asp Arg Val
 1000 1005 1010 1015
 GCT GAG GTT CAG CAG GTA CTG CGG CCA GCA GAA AAG CTG GTG ACA AGC 3846
 Ala Glu Val Gln Gln Val Leu Arg Pro Ala Glu Lys Leu Val Thr Ser
 1020 1025 1030
 ATG ACC AAG CAG CTG GGT GAC TTC TGG ACA CGG ATG GAG GAG CTC CGC 3894
 Met Thr Lys Gln Leu Gly Asp Phe Trp Thr Arg Met Glu Glu Leu Arg
 1035 1040 1045
 CAC CAA GCC CGG CAG CAG GGG GCA GAG GCA GTC CAG GCC CAG CAG CTT 3942
 His Gln Ala Arg Gln Gln Gly Ala Glu Ala Val Gln Ala Gln Gln Leu
 1050 1055 1060
 GCG GAA GGT GCC AGC GAG CAG GCA TTG AGT GCC CAA GAG GGA TTT GAG 3990
 Ala Glu Gly Ala Ser Glu Gln Ala Leu Ser Ala Gln Glu Gly Phe Glu
 1065 1070 1075
 AGA ATA AAA CAA AAG TAT GCT GAG TTG AAG GAC CGG TTG GGT CAG AGT 4038
 Arg Ile Lys Gln Lys Tyr Ala Glu Leu Lys Asp Arg Leu Gly Gln Ser
 1080 1085 1090 1095
 TCC ATG CTG GGT GAG CAG GGT GCC CGG ATC CAG AGT GTG AAG ACA GAG 4086
 Ser Met Leu Gly Glu Gln Gly Ala Arg Ile Gln Ser Val Lys Thr Glu
 1100 1105 1110
 GCA GAG GAG CTG TTT GGG GAG ACC ATG GAG ATG ATG GAC AGG ATG AAA 4134

112

GAC ATG GAG TTG GAG CTG CTG CGG GGC AGC CAG GCC ATC ATG CTG CGC 4182
Asp Met Glu Leu Glu Leu Leu Arg Gly Ser Gln Ala Ile Met Leu Arg
1130 1135 1140
TCA GCG GAC CTG ACA GGA CTG GAG AAG CGT GTG GAG CAG ATC CGT GAC 4230
Ser Ala Asp Leu Thr Gly Leu Glu Lys Arg Val Glu Gln Ile Arg Asp
1145 1150 1155
CAC ATC AAT GGG CGC GTG CTC TAC TAT GCC ACC TGC AAG T 4270
His Ile Asn Gly Arg Val Leu Tyr Tyr Ala Thr Cys Lys
1160 1165 1170
GATGCTACAG CTTCCAGCCC GTTGCCCCAC TCATCTGCCG CCTTTGCTTT TGGTTGGGGG 4330
CAGATTGGGT TGAATGCTT TCCATCTCCA GGAGACTTTC ATGCAGCCTA AAGTACACCC 4390
TGGACCACCC CTGGTGTGTA GCTAGTAAGA TTACCCTGAG CTGCAGCTGA GCCTGAGCCA 4450
ATGGGACAGT TACACTTGAC AGACAAAGAT GGTGGAGATT GGCATGCCAT TGAAACTAAG 4510
AGCTCTCAAG TCAAGGAAGC TGGGCTGGGC AGTATCCCCC GCCTTTAGTT CTCCACTGGG 4570
GAGGAATCCT GGACCAAGCA CAAAACTTA AAAAAAGTGA TGTA AAAATG AAAAGCCAAA 4630
TAAAAATCTT TGGAAAAGAG CCTGGAGGTT CAACGAG 4667

Sequence No.: 26

Sequence length: 1086

Sequence type: Nucleic acid

Strandedness: Double

Topology: Linear

Sequence kind: cDNA to mRNA

Original source:

Organism species: *Homo sapiens*

Cell kind: Stomach cancer

Clone name: HP10298

113

Existence site: 138.. 506

Characterization method: E

Sequence description

TTTAATTTCC CCGAAATCAG ACTGCTGCCT TGGACCGGGA CAGCTCGCGG CCCCCGAGAG	60
CTCTAGCCGT CGAGGAGCTG CCTGGGGACG TTTGCCCTGG GGCCCCAGCC TGGCCCCGGT	120
CACCCTGGCA TGAGGAG ATG GGC CTG TTG CTC CTG GTC CCA TTG CTC CTG	170
Met Gly Val Leu Leu Leu Val Pro Leu Leu Leu	
1 5 10	
CTG CCC GGC TCC TAC GGA CTG CCC TTC TAC AAC GGC TTC TAC TAC TCC	218
Leu Pro Gly Ser Tyr Gly Leu Pro Phe Tyr Asn Gly Phe Tyr Tyr Ser	
15 20 25	
AAC AGC GCC AAC GAC CAG AAC CTA GGC AAC GGT CAT GGC AAA GAC CTC	266
Asn Ser Ala Asn Asp Gln Asn Leu Gly Asn Gly His Gly Lys Asp Leu	
30 35 40	
CTT AAT GGA GTG AAG CTG GTG GTG GAG ACA CCC GAG GAG ACC CTG TTC	314
Leu Asn Gly Val Lys Leu Val Val Glu Thr Pro Glu Glu Thr Leu Phe	
45 50 55	
ACC CGC ATC CTA ACT GTG GGC CCC CAG AGC CTG GGG TCC GAA GCT TTG	362
Thr Arg Ile Leu Thr Val Gly Pro Gln Ser Leu Gly Ser Glu Ala Leu	
60 65 70 75	
GCT TCC CCG ACC CGC AGA GCC GCT TGT ACG GTG TTT ACT GCT ACC GCC	410
Ala Ser Pro Thr Arg Arg Ala Ala Cys Thr Val Phe Thr Ala Thr Ala	
80 85 90	
AGC ACT AGG ACC TGG GGC CCT CCC CTG CCG CAT TCC CTC ACT GGC TGT	458
Ser Thr Arg Thr Trp Gly Pro Pro Leu Pro His Ser Leu Thr Gly Cys	
95 100 105	
GTA TTT ATT GAG TGG TTC GTT TTC CCT TGT GGG TTG GAG CCA TTT	503

TAAGTGT TTTTATACTT CTCAATTAA ATTTTCTTTA AACATTTTT TACTATTTTT 560
TGTAAGCAA ACAGAACCCA ATGCCTCCCT TTGCTCCTGG ATGCCCCACT CCAGGAATCA 620
TGCTTGCTCC CCTGGGCCAT TTGGGGTTTT GTGGGCTTCT GGAGGGTTCC CCGCCATCCA 680
GGCTGGTCTC CCTCCCTTAA GGAGGTTGGT GCCCAGAGTG GCGGGTGGCC TGTCTAGAAT 740
GGCGCCGGA GTCCGGGCAT GGTGGGCACA GTTCTCCCTG CCCCTCAGCC TGGGGGAAGA 800
AGAGGGCCTC GGGGGCCTCC GGAGCTGGGC TTTGGGCCTC TCCTGCCCAC CTCTACTTCT 860
CTCTCAGGC GCTGACCCA GTCTGCCCAC TGAGGGGCTA GGGCTGGAAG CCAGTTCTAG 920
GCTTCCAGGC GAAAGCTGAG GGAAGGAAGA AACTCCCCTC CCCGTTCCCC TTCCCCTCTC 980
GGTTCCAAAG AATCTGTTTT GTTGTCATTT GTTTCTCCTG TTCCCTGTG TGGGGAGGGG 1040
CCCTCAGGTG TGTGTACTTT GGACAATAAA TGGTGCTATG ACTGCC 1086

Sequence No.: 27

Sequence length: 866

Sequence type: Nucleic acid

Strandedness: Double

Topology: Linear

Sequence kind: cDNA to mRNA

Original source:

Organism species: *Homo sapiens*

Cell kind: Stomach cancer

Clone name: HP10368

Sequence characteristics:

Code representing characteristics: CDS

Existence site: 73.. 600

Characterization method: E

Sequence description

ACTCAGAAGC TTGGACCGCA TCCTAGCCGC CGACTCACAC AAGGCAGGTG GGTGAGGAAA 60

115

1	5	10	
GCC CTC TCC TAC ACT CTG GCC AGA GAT ACC ACA GTC AAA CCT GGA GCC			159
Ala Leu Ser Tyr Thr Leu Ala Arg Asp Thr Thr Val Lys Pro Gly Ala			
15	20	25	
AAA AAG GAC ACA AAG GAC TCT CGA CCC AAA CTG CCC CAG ACC CTC TCC			207
Lys Lys Asp Thr Lys Asp Ser Arg Pro Lys Leu Pro Gln Thr Leu Ser			
30	35	40	45
AGA GGT TGG GGT GAC CAA CTC ATC TGG ACT CAG ACA TAT GAA GAA GCT			255
Arg Gly Trp Gly Asp Gln Leu Ile Trp Thr Gln Thr Tyr Glu Glu Ala			
50	55	60	
CTA TAT AAA TCC AAG ACA AGC AAC AAA CCC TTG ATG ATT ATT CAT CAC			303
Leu Tyr Lys Ser Lys Thr Ser Asn Lys Pro Leu Met Ile Ile His His			
65	70	75	
TTG GAT GAG TGC CCA CAC AGT CAA GCT TTA AAG AAA GTG TTT GCT GAA			351
Leu Asp Glu Cys Pro His Ser Gln Ala Leu Lys Lys Val Phe Ala Glu			
80	85	90	
AAT AAA GAA ATC CAG AAA TTG GCA GAG CAG TTT GTC CTC CTC AAT CTG			399
Asn Lys Glu Ile Gln Lys Leu Ala Glu Gln Phe Val Leu Leu Asn Leu			
95	100	105	
GTT TAT GAA ACA ACT GAC AAA CAC CTT TCT CCT GAT GGC CAG TAT GTC			447
Val Tyr Glu Thr Thr Asp Lys His Leu Ser Pro Asp Gly Gln Tyr Val			
110	115	120	125
CCC AGG ATT ATG TTT GTT GAC CCA TCT CTG ACA GTT AGA GCC GAT ATC			495
Pro Arg Ile Met Phe Val Asp Pro Ser Leu Thr Val Arg Ala Asp Ile			
130	135	140	
ACT GGA AGA TAT TCA AAC CGT CTC TAT GCT TAC GAA CCT GCA GAT ACA			543
Thr Gly Arg Tyr Ser Asn Arg Leu Tyr Ala Tyr Glu Pro Ala Asp Thr			

CTC TCC TAC ACT CTG GCC AGA GAT ACC ACA GTC AAA CCT GGA GCC
 Ala Leu Ser Tyr Thr Leu Ala Arg Asp Thr Thr Val Lys Pro Gly Ala

116

Ala Leu Leu Leu Asp Asn Met Lys Lys Ala Leu Lys Leu Leu Lys Thr

160

165

170

GAA TTG TAAAGAAAAA AAATCTCCAA GCCCTTCTGT CTGTCAGGCC TTG

640

Glu Leu

175

AGACTTGAAA CCAGAAGAAG TGTGAGAAGA CTGGCTAGTG TGGAAGCATA GTGAACACAC

700

TGATTAGGTT ATGGTTTAAT GTTACAACAA CTATTTTTTA AGAAAAACAA GTTTTAGAAA

760

TTTGGTTTCA AGTGTACATG TGTGAAAACA ATATTGTATA CTACCATAGT GAGCCATGAT

820

TTTCTAAAAA AAAAAATAAA TGTGTTGGGG GTGTTCTGTT TTCTCC

866